Ministero dell' Istruzione,
dell' Università e della Ricerca

Dipartimento di Scienze
Fisiche ed Astronomiche

Università degli Studi
di Palermo

UNIVERSITÁ DEGLI STUDI DI PALERMO

Dipartimento di Scienze Fisiche ed Astronomiche

Dottorato di Ricerca in Fisica – XXII ciclo

S.S.D.: FIS/07

Intermolecular interactions in metastable protein solutions

Palermo, Febbraio 2011

Coordinatore: Prof. Antonio Cupane

Supervisore: Prof. Maurizio Leone

Candidato: Samuele Raccosta

Co-supervisore: Dr. Mauro Manno

To my parents

List of abbreviations and symbols

CD	Circular Dichroism
PL	Photoluminescence
MRE (θ)	Mean Residue Ellipticity
HEWL	Hen Egg-White Lysozyme
α-Sn	α-synuclein
DSC	Differential Scanning Calorimetry
AFM	Atomic Force Microscopy
TFE	2-2-2 Thrifluoroethanol
ThT	Thioflavin T
SAXS	Small Angle X-ray Scattering
SALS	Small Angle Light Scattering
DLS	Dynamic Light Scattering
ΔG_{NU}	Change of free energy in the native-unfolded transition
ΔH	Change of enthalpy
Hb	Hemoglobin
SANS	Small Angle Neutron Scattering
B_2	Second virial coefficient
h_2	Hydrodynamic coefficient
D_c, D_s	Collective or Self Diffusion coefficient
NLPB	Non Linear Poisson Boltzmann equation
FTIR	Fourier Transform InfraRed spectroscopy
NAG	N-acetyl-glucosamine
NAM	N-acetyl-muramic acid
L-E	Lumry-Eyring
k_B	Boltzmann constant
c_p^{exc}	excess of specific heat
PBS	Phosphate Buffer Solution
OD	Optical Density
UV-VIS	Ultra-Violet Visible
R_G, R_g, R_h, R_i	Guinier/Gyration/Hydrodynamic/Interaction Radius
TRP	Tryptophan

DTT	Dithiothreithol
GuHCl	Guanidinium Hydrochloride Acid
GFP	Green Fluorescence Protein
SLS	Static Light Scattering
PMF	Potential of Mean Force
H-S	Hard Sphere
N_A	Avogadro's number
I	Ionic strength
APBS	Adaptive Poisson-Boltzmann Solver
NAC	Non-amyloid β component
SVD	Singular Value Decomposition

Contents

Summary

The aggregation processes of proteins are very useful from biotechnological point of view, although sometimes they are pathological. Proteins are very complex objects and, in suitable conditions (typically the physiological ones), they can develop very complex functions in living beings. This variety and complexity in the structural conformation and biological function of proteins is mirrored by a great variety of aggregation processes and morphologies, such as amorphous particulate, fibrils, crystals, gel, etc. Often, proteins with different physical and chemical characteristics, lead to the same aggregation process in different conditions. Therefore the details of a protein are not decisive for the evolution of aggregation, although they can introduce some differences.

Each protein, in certain conditions, interacts with other proteins by an intermolecular potential, characterized by different contributions. Someones are enthalpic-like, others are entropic-like and all terms are different in their range of action. By changing experimental conditions, these contributions change, however sometimes the aggregation process remains unchanged in its general features. This suggests that the balance of the different contributions is very important, more than the characteristics of the individual contributions. A similar mechanism was highlighted in crystallization [87], but it is not clear if it can be extended to other aggregation processes.

In order to obtain new insights about that, we could study several proteins in the same conditions or one single protein in different conditions. We chose the second strategy for convenience and we focused our study on a model system, hen egg-white lysozyme, a very known protein. That allowed us to not waste any time in searching for details and to focus our attention on more general and interesting questions, as the balance of different energetic contributions in certain conditions. Some of these contributions are attractive and induce the aggregation; others are repulsive and stabilize the solution. We want to study the different aggregation pathways, so we need to enhance the attractive contribution of intermolecular potential. The most simple way to do that is the increasing of the temperature. Indeed high temperatures

induce the partial unfolding of proteins, the exposure of hydrophobic regions and the acceleration of the self-assembly reaction. We started our investigations by studying the aggregation process and the formed structures in the most typical biological condition, that is at physiological pH.

Also, we evaluated the effect of other physical and chemical parameters on the aggregation pathway. In particular we added trifluoroethanol into the solution. As all alcohols, it changes the dielectric constant of the solvent and increases the electrostatic repulsive contribution. On the other hand, TFE is known to affect the protein conformation by stabilizing α-structure. As we highlighted, TFE stabilizes an unfolding intermediate state, which reduces the hydrophobic effect. Therefore TFE can affect the balance of the interaction contributions in different ways.

Also, we investigated the role of electric charge on the interaction potential. By decreasing pH, the net charge of lysozyme strongly increases and the electrostatic repulsion acts both at intramolecular level by modifying and destabilizing the protein structure and at intermolecular level by stabilizing the solution. We studied the effect of the high temperature on the unfolding as an enhancer of the aggregation process. It changes the balance between attractive and repulsive contributions and can strongly affect the aggregation pathway or its evolution rate. We also modulated the range of action of electrostatic interaction by ionic strength, which typically affects the speed of reaction. In the amorphous aggregation this could not have a relevant role, but in an ordered aggregation process, the increasing of the speed of reaction could drive the system on another pathway. Therefore ionic strength allows us to rule finely the repulsive contribution without affect sensitively the others.

Although we studied the balance of the interactions which drive the different aggregation pathways, our investigations are focused on fibrillation. We highlighted the presence of an intermediate state in TFE solutions and conformational changes at acidic pH: both events are correlated with the fibril formation, so a structural change seem necessary for this pathway. However it is not sufficient, because amorphous aggregation is also started by conformational change. We have evidences of formation of oligomers at acidic pH and this suggests that the structural changes are specific because the monomers are assembled into oligomers in a reproducible way. We do not have the same evidences for lysozyme in TFE solutions yet, but it would be interesting to clear, as a future project, if oligomers are also formed in these conditions.

Therefore very useful informations could come from a structural characterization of the oligomers. A few studies were already led in this purpose and, although some morphologies seem preferred (spherical or anular

oligomers), the situation is not clear. We studied the oligomeric species of a mutant α-synuclein to understand if the informations and the conclusions obtained for lysozyme oligomers are still valid. Finally we tried to explain the physical reasons which drive the oligomer formation and the fibril elongation.

Chapter 1

The argument: thermodynamic and conformational stability of proteins

1.1 Proteins: marginally stable objects

The biological and biochemical processes in living beings are based on proteins. These are polipeptidic chains selected by nature to perform several functions. Their functionality is guaranteed by their flexibility, that is their capability to modulate their structural conformation [79]. Some proteins do not need a high flexibility as e.g. those having a transport function or able to catalyze simple reactions. Other most complex proteins can perform large translocation processes or work as molecular motors [109]. In the framework of thermodynamics, these changes are often described in terms of *energy landscape*: it represents the energy of the system in the phase space of coordinates and momenta [80]. For a simple visualization, the energy landscape is often proiected upon some effective coordinates: conformational and reaction coordinates (Fig. 1.1). The conformation coordinates indicate the changes at the single molecule level regarding both secondary and tertiary structure. The reaction coordinates may be used to indicate the pathway of *self-assembly*, which depends upon molecular conformation. It is selected by the environmental conditions and the intermolecular interactions involved. In particular, one could trace some paths, along the conformational coordinates, from regions where proteins are chain-like to region where they are in the functional final state (*native*). This process is called *folding* and several theories were suggested about that [90]. In given environmental and thermodynamic conditions, proteins do not have a unique conformation, but

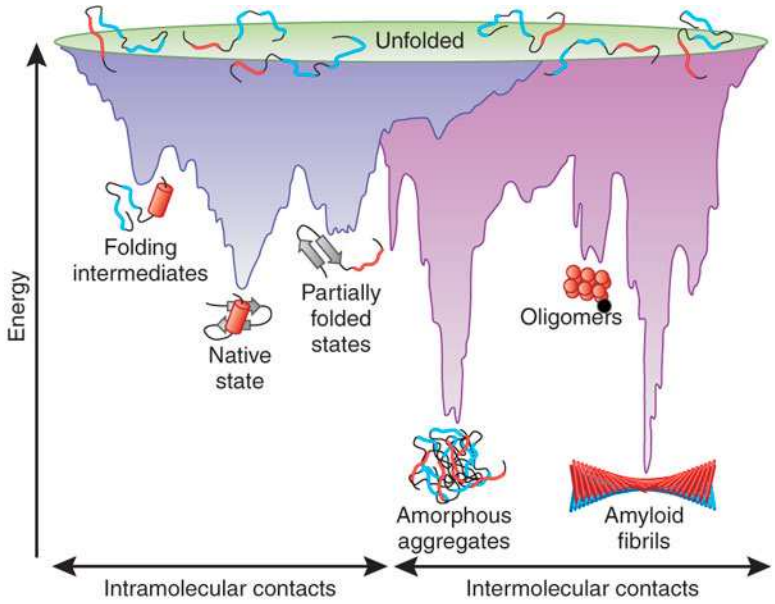

Fig. 1.1: Schematic representation of energy landscape of a protein. The enthalpy is represented on the vertical axis. The entropy is represented in the orizontal plane and distinguish among different isoenergetic states. Two different origins of the entropic contribution are highlighted: the variability at intramolecular level or *conformational coordinate* (blue part) and the variability at intermolecular level or *reaction coordinate* (violet part). Adapted from [106]

it fluctuates among quite isoenergetic equilibrium conformations (*substates*) [79]. However some proteins can be statistically trapped in a local minimum, and this is the origin of the intrinsic instability of a protein solution after long times. How much stable a protein is? Calorimetric investigations highlighted that the free energy difference between the native and unfolded state, ΔG_{NU}, lies in the range -5 ÷ -15 $kcal/mol$ [138]. It is a quite small value with respect to k_BT (≈ 0.59 $kcal/mol$) and the strength of hydrogen bond in bulk water (≈ 4.3 $kcal/mol$). This is the reason why protein stability can be dramatically affected by small environmental changes or by mutation of only one residue. Protein stability is due to different types of contributions. When a protein folds, the number of accessible spatial configurations is reduced, so that chain conformational entropy decreases. This would be an unfavourable contribution to the stability (since $-T\Delta S_{conf}>0$). A favourable contribution arises from the hydrophobic effect which is an entropic term, and from the van der Waals forces, which give an enthalpic contribution. Salt bridges and hydrogen bonds have an ambivalent character in protein stabilization. It is very fascinating that a so large amount of different type of interactions can

be almost completely balanced and allow the protein to be enough flexible and perform its biological function. Therefore, proteins can be considered as *marginally stable* in a thermodynamic framework [248]. This is at the origin of several conformational diseases (e.g. HbS [178]).

1.2 Protein stability in stress conditions

As we highlighted in the last paragraph, the stability of proteins is very mild because the difference in the free energy between the native and the unfolded state is small compared to typical k_BT values. Stress conditions in the environment can further reduce this energy difference and make more simple the transition to a not native state.

The pH of the solution has a relevant role in the protein stability. Each protein works with a variable efficiency within a certain range of pH [167]. The protonation equilibrium of the lateral residues of amminoacids determines the net charge of the protein, the intramolecular electrostatic interactions and can modify its structure in a not physiological pH.

Crowding is another important factor that can affect stability of proteins and their dynamics. It is a central step in the knowledge of the behavior of a protein *in vivo*. Indeed crowding is the typical condition of the intracellular environment, where a high concentration of different proteins and metabolites is present (up to 350 mg/ml in the cytoplasm [182]). These conditions reduce the number of accessible structural configurations and the dynamics of proteins can be quite different from that observed in dilutes solutions. This can be observed by using a huge range of concentrations or by embedding the protein in a matrix [49, 224] .

Protein stability is strongly conditioned by temperature. Heating speeds up the protein dynamics and the reactions, but if the temperature is too high not native states could be stabilized with respect to the native one. Temperature enhances the role of entropic contributions to the protein thermodynamics. On the other hand, the freezing slowdowns dynamics, reactions and the solvent is tipically more structured. In case of the water as a solvent, the formation of ice-like structures can have similar effects than crowding [247].

1.3 Instability and self-assembly

Thermodynamic instability typically drives proteins towards the self-assembly into supramolecular structures. The study of self-assembly is very important in several fields of research. Some types of aggregation processes form

dangerous structures: *amyloid plaques* in a degenerated brain (Alzheimer's, Parkinson's, Creutzfeld-Jacob desease, spongiform encefalopaties and others [201, 43]), HbS polymers in falciform anemia, condensed γ-crystallin in cataract and so on. Other aggregation processes are useful for different purposes. For example, an ordered protein self-assembly, *crystallization*, is useful for protein purification and identification of the tertiary structure. As a further example, the formation of *gels* and networks of aggregated proteins [45] can be used in food technologies (to change the texture of foods [77]), cosmetics and as a surface for cellular *scaffold* [269].

The study of protein instability can be addressed by two main approaches. The first is the conformational one. Slight variations in physical and chemical conditions of solution or the addition of cofactors can modify the equilibrium among conformational substates and provide enough energy to induce a relevant change in protein structure. The identification of particular motifs or the composition of the secondary structure gives informations about the behavior of the protein in solution. In general, different partially unfolded conformations can discriminate among different aggregation morphologies [43]. However, different conformational states can induce the same type of aggregation, although the aggregation rate depends from the particular intermediate [69]. In the formation of not-ordered, amorphous, aggregates, as suggested by the name, the specificity of the contacts is typically less important than in the formation of regular structures like crystals or fibrils. Thus amorphous aggregates may easily form with no specific intermediate conformation. Typically the presence of β-sheet in the secondary structure is a sign of tendency to aggregate and is at the origin of the formation of peculiar elongated aggregates called amyloid fibrils.

A very useful equation which summarizes this approach is the Chiti-Dobson equation [61, 258]. It is a phenomenological expression which can be used to make a prediction of aggregation rate of a polypeptide chain. The rate of aggregation is related to intrinsic (protein dependent) and extrinsic (condition dependent) factors. In particular, it depends upon the hydrophobicity of the sequence, the net charge, the pH of the solution, the ionic strength and the polypeptide concentration. Since the protein solutions are quite stable in physiological conditions, it is reasonable that a symmetry breaking at molecular level (conformational change) leads to a simmetry breakage at supermolecular level (aggregation). This approach is mainly used to study neurodegenerative pathologies and other deseases which are related to conformational changes of protein structure (e.g. prions [201]) or to changes of aminoacidic sequence, like in mutants [249, 43]. For example, HbS is responsible of sickle cell anemia because of the substitution of two hydrophilic amino acids by two hydrophobic ones [71].

The second approach used to study the instability may be called the thermodynamic one. It is focused on the identification of intermolecular interactions relevant for a protein. They have different physical origins and a different range of action. One of the main contributions is the electrostatic one: it is a long range interaction, eventually screened by electrolytes, and has enthalpic origins. Another main contribution is the hydrophobic effect due its entropic origin and involves the whole system. A lot of other interactions (hydrogen bonds, van der Waals forces, etc) and solvent mediated effects rule the behavior of the protein system. In particular conditions, the delicate balance of intra- and intermolecular interactions can prevent the protein aggregation and stabilize the solution. When this balance is broken, by changing some physical or chemical parameters, the instability of solution leads to self-assembly of the proteins. However the pathway of aggregation is not unique, but it depends upon which are the most relevant contributions and their relative weight. In the case of crystallization, proteins can assemble into an ordered arrangement. At the microscopical level, the ordered assembly has been proposed to rise from the balance of aggregation rate, dissociation rate and the probability of internal reordering [54]. If the attractive contribution has a too short range, the particle assembles in amorphous configuration. However, if there is enough time between two consecutive attachments, the particle can rearrange its position because the crystalline configuration is more stable [54]. Another ordered process of aggregation is fibrillation, that is the formation of elongated structures. Sometimes fibrils are treated as linear crystals because have translational invariance and the physics of these two processes has some elements in common. For example, they are both nucleated processes. In case of crystals the formation of nucleus can be a one step or a two steps process. In the first case the process can be described classically in terms of an energy barrier. In the second case a not classical theory involving fluctuations of concentration is necessary [84]. In particular the nucleus would form within a metastable droplet at higher protein concentration than bulk [174, 100, 237]. On the other hand, the thermodynamics of formation of fibrillar elongated structures and their precursors is not definitively clear yet. In fact HbS forms fibrils by stacking of single monomers, but other types of fibrils seem to be constituted by the addition of preformed oligomers.

Fluctuations of concentration could have a relevant role. These were observed through SAXS and SANS measurements on highly concentrated neutral solutions of lysozyme and lead to a local clusterization of proteins, although a heated debate between Schurtenberger's and Svergun's groups is running about the life-time of clusters [236, 233]. A clusterization could favour the binding contacts. Also the balance of long range electrostatic

repulsion and short range colloid-like attraction seems important for the formation of elongated and spiral-like clusters (*spirals of Bernal*) [228].

1.4 The present work: protein stability and different pathways of self-assembly

The general aim of this thesis, is the understanding of role which a different balance among the intermolecular interactions of proteins has on different aggregation pathways, with a special attention for fibrillogenesis, interesting from a medical point of view. We used Hen Egg-White Lysozyme (HEWL) as a model system because it has been studied from different points of view since several years. We used this knowledge as a background to focus our attention on some, still puzzling, details. Since we want to study the different aggregation pathways, we need to enhance the attractive contribution of intermolecular potential. The most simple way to do that is to increase the temperature, in order to exploit the increased hydrophobic effect due to the partial thermal unfolding of the protein. Afterwards we modulate the intermolecular interactions by adding alcohols (which modify the properties of the solvent) or changing the electrostatic repulsion (by varying the protein charge or the ionic strength). Several works address the role of these physical and chemical parameters on protein solutions, but the originality of our work, in addition to the obtained results, lies in the interpretation about their role on fibrillogenesis.

Chapter 1 introduces the topic of this thesis, that is the thermodynamic and conformational stability of proteins and the aggregation pathways taken by a different balance of interactions.

Chapter 2 is divided into two sections and addresses the study on lysozyme stability at physiological pH by focusing the attention on the role of temperature. At room temperature lysozyme solutions are quasi-stable. When temperature grows, the entropic contributions increase their weight with respect to the electrostatic ones. Typically the unfolding is considered as a two-state equilibrium transition, but this is only an approximation. In the first section we elicit the effects of aggregation upon the unfolding and upon the reversibility of the native-unfolded state transition. On the other hand we study the effects of the amount of unfolding on the aggregation by analizing self-assembly kinetics within the temperature range around the unfolding temperature. A quantitative study leads to establish a cause-effect relation between unfolding and aggregation.

Studying the pathway of aggregation at temperatures where the interme-

diate states are very populated and characterized by a different entropic contribution for each temperature and, eventually, for each protein, is very hard. Therefore, in the second section we focus on the mechanism of the aggregation of lysozyme at temperature enough high to guarantee for the conformational equilibrium of the protein is shifted towards its thermally denatured state. We study this process as a function of the protein concentration. We investigate the structure of aggregates and elicit the mechanical properties of these samples through rheological characterization. This approach allows us to catch other informations about intermolecular interactions which stabilize the aggregates. A sol-gel transition emerges from experiments.

In the Chapter 3 we study more deeply the role of TFE upon the stability of lysozyme. TFE is an alcohol well-known to affect the protein conformation. In addition to the general interest of this topic, a recent work suggests that lysozyme, in these conditions and at high temperatures, forms fibrils [130]. We tried to focus on the mechanism of action of TFE. Steady state and time resolved photoluminescence of triptophans revealed changes in the polarity of the environment around tryptophans exposed to the solvent and then around the overall protein, caused by preferential solvation of lysozyme. We elicit that it is able to change the protein conformation. Besides the study of thermal unfolding of lysozyme at different TFE concentration allows us to elicit the role of alcohol on protein stability. Eventually we investigate the morphology of aggregates formed at quite high concentration of lysozyme and different TFE concentration through AFM imaging, ThT assays and eye visualization to have some confirmation about lysozyme fibrillation in these conditions.

Lysozyme can form fibrils also at very acidic pH and in the Chapter 4 we study the mechanism of fibrillogenesis at pH 2 to have new insights on the role of electric charge on this pathway. Indeed in this case the origin of fibrillation and oligomer formation is not due to a co-solvent, but to the combined effect of strong repulsive interactions and high entropic contribution enhanced by increasing the temperature. Lysozyme solutions at acidic pH in the low concentration regime are studied to avoid every aggregation effects. We focus on the immediate effects of the increase of temperature on thermodynamic stability of lysozyme and on its conformation. Afterwards we investigate the intermolecular interactions which characterize our system in these conditions, at the onset of incubation. We measured osmotic compressibility by light scattering and derived the second virial coefficient, B_2, which is an indicator of the intermolecular interactions and of the solvent quality [74]. We find that the value of B_2 can be largely explained by considering excluded volume and electrostatic repulsion. Electrostatics is discussed in the context of DLVO theory and also by taking into account the Manning condensation theory.

We solve the Non Linear Poisson-Boltzmann equation, to have a model-free prediction of the dependence of the electrostatic potential upon changes of ionic strength, pH and temperature. After this characterization of the involved interactions, we study the fibrillogenesis at higher concentrations (about 18 mg/ml) at three temperatures in the range 60–70 $°C$ by focusing on the individuation of the species involved in the process. We find the formation of oligomers at high temperature. Their hydrodynamic radius is about 10 nm and are formed during early hours of *lag-phase* through a saturation exponential behavior. This result is very important because the detection of oligomers is not simple, because of their metastability.

In order to directly observe oligomers and fibrils one needs to stabilize them and use AFM or electron microscopy. In this case some treatment of the sample is necessary and could affect the morphology of the observed objects. On the contrary, we obtain the kinetics of formation of oligomers in the same experimental conditions for fibril formation, in a not invasive way. Finally we highlight, through AFM imaging, the morphological and structural characteristics of fibrils, obtained after several days of incubation.

The formation of oligomeric species during the fibrillogenesis is investigated in Chapter 5 through an advanced analysis of SAXS form factors on the mutant α-Synuclein A30P. We used this protein because some studies on determination of the structure of lysozyme oligomers were already recently performed by electronic microscopy and FTIR [78]. Also we were interested to understand if our conclusions about lysozyme could be extended to another protein. We have not SAXS data about lysozyme yet, but it could be a hint for future works. Lysozyme and α-synuclein have similar molecular mass, but very different secondary and tertiary structures and are both involved in degenerative deseases: lysozyme is related to hereditary non-neuropathic systemic amyloidosis and α-synuclein is related to Parkinson's [235]. This section shows a method to "isolate" the form factor of the oligomer from data and, eventually, build a $3D$ model. SAXS form factors were collected at the beamline X33 in Hamburg as a function of the time during fibrillogenesis kinetics. We compare kinetics for the wild type [89] and the mutant. The results obtained so far are not exhaustive of the structure formed in the fibrillogenesis process. However, they clearly point toward the presence of a variety of morphologies and they encourage us to continue the work by taking into account the structural heterogeneity.

Chapter 6 contains a summary of main results and the final remarks.

Appendix A contains the description of the used equipments, the experimental procedures and some tests of reliability.

Appendix B contains a summary of the didattical and research activities developed during the PhD course.

Chapter 2

Lysozyme stability at physiological pH

2.1 The model system: HEWL

In order to achieve the aim of this thesis (see § 1.4) we decided to study one single protein on different conditions. We chose, as a model system, Hen Egg-White Lysozyme (HEWL), a well-known protein. That allowed us to focus our attention only on more general and interesting points.

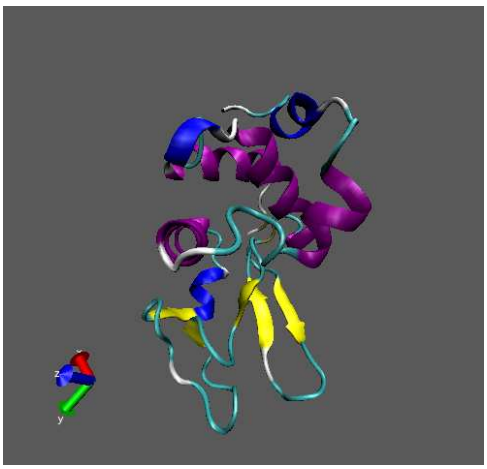

Fig. 2.1: Tertiary structure of hen egg-white lysozyme. From Protein Data Bank AKI1

Lysozyme is an hydrolytic enzyme and its biological activity is to break the 1-4 β-glicosidic bonds between *N-acetyl-glucosamine* (NAG) and *N-acetyl-muramic acid* (NAM) of bacterial membrane. For this reason it is present

in several biological fluids (milk, saliva, tears) and organs (liver, kidney) as aspecific immunitary defence and it is used as antibacteric additive in food industry [159]. Lysozyme is a globular protein and can be approximately modelled by a prolate ellipsoid with size 45x26x26 $\mathring{A}^3$. It is made of 129 amino acids organized in an α/β-structure (Fig. A.6). In particular, the tertiary structure is made of two domains: the former contains four α-helices and an helix α_{3-10}; the second domain contains an helix α_{3-10} and an antiparallel β-sheet made of three filaments. These domains are connected by an antiparallel β-sheet made of two filaments [57]. Lysozyme has four disulfide bridges and no free cysteines. Its molecular mass is 14.3 kDa and it has an isoelectric point of 11.2 [246]. It is a quite heat stable protein and manteins its functionality in a huge range of pH. Lysozyme can form different aggregated morphologies.

2.2 Lysozyme aggregation induced by thermal unfolding

2.2.1 Introduction

As we highlighted in § 1.4, we want to study the different aggregation pathways, so we need to enhance the attractive term of intermolecular potential with respect to others. The most simple way to do that is the increase of the temperature. When we boil an egg the egg-white, very rich of proteins, changes color and texture. These macroscopical observations reflect the modifications of the state of protein. The rise of temperature can alter the energy landscape of a protein system in a specific or not specific way. In the first case, the protein can change the arrangement of a restricted area because the stability of this part is less than the rest [11]. In fact several experimental works demonstrate that the increase of temperature can induce a conversion from α to β-structure [97, 158] or some similar internal conversion of the secondary structure. In the second case, modifications can be due to the breaking of some hydrogen bonds and salt bridges; this leads to a global destabilization of the protein structure and some swelling [225]. In both cases, hydrophobic residues are exposed to the solvent. Hydrophobic effect has entropic origins and intrinsecally depends on the polarity of the solvent where protein is [51]. Hydrophobic regions tend to assembly each other to minimize the surface exposed to the polar solvent. This attractive contribution can be relevant with respect to the electrostatic repulsion or be overwhelming. This is a key concept in the physics of protein aggregation. The rise of temperature can populate some partially unfolded conformations (*intermediates*) and these

have a very important role in the study of the folding pathway and in the aggregation processes, as we discussed in Chapter 1.

We studied the effects of aggregation upon the unfolding and upon the reversibility of this transition. On the other hand we focused on the effects of unfolding degree on the self-assembly by studying the kinetics at different temperatures in proximity of the unfolding temperature. Besides we put quantitatively on relationship the unfolding enthalpy and the activation energy of the aggregation process and studied the role of ionic strength on the unfolding-aggregation process.

2.2.2 DSC experiments

Two consecutive thermograms were run at 0.13 $°C/min$ on the same sample (Fig. 2.2). The first ramp shows the endothermic peak of unfolding, while the second one is flat. Therefore there is not a net exchange of heat after the first upward scan and the sample results thermodynamically irreversible. At higher scan rate (20 $°C/min$) the second ramp (triangle down) exhibits a similar profile to former one but it is reduced. Theoretically, we may consider

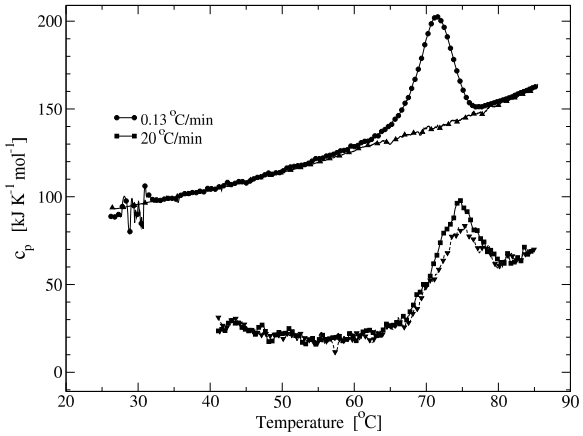

Fig. 2.2: Specific heat c_p (not buffer subtracted) vs temperature. Consecutive DSC upward ramps at two different scan rates. 0.13 $°C/min$: first scan (circle-line), second scan (triangle up-line); 20 $°C/min$: first scan (square-line), second scan (triangle down-line)

two possible explanations of irreversibility. It could be caused by irreversible unfolding of single protein [16, 17] or by intermolecular aggregation [208]. If only the first hypothesis were right, when the sample comes back to the

room temperature the protein would make its conformation more compact and would establish new intramolecular bonds, even if they are not native. Therefore the second upward scan should not be flat, but it should measure the cleavage of these new bonds. In this case we would not necessarily expect the same thermogram, but it cannot be flat. That allow us to argue that the irreversibility is mainly due to a slow process like intermolecular aggregation, characterized by a time scale compatible with the range of the highest scan rates used in DSC measures [65].

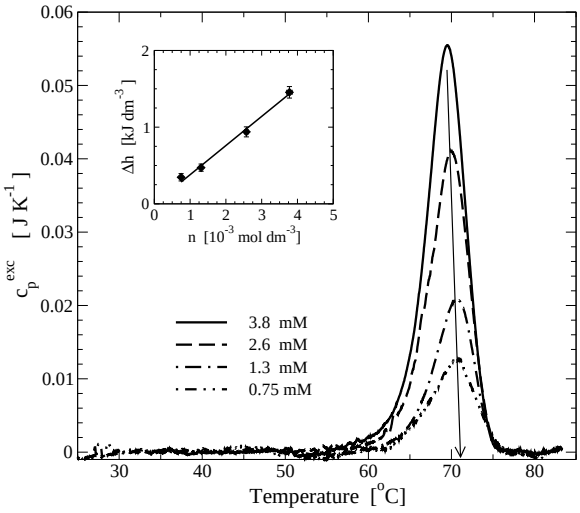

Fig. 2.3: Excess of specific heat at different lysozyme concentrations and 0.13 $°C/min$. The arrow is a guide for the eyes through the peak maxima and towards the extrapolation at infinite dilution $T_{1/2}$=71.1 $°C$. Inset: enthalpy density vs molar protein concentration. The solid line is a linear fit.

We performed some DSC measurements in function of protein concentration to elicit the effect of aggregation on thermodynamic parameters (Fig. 2.3). Inset of Fig. 2.3 shows that the enthalpy density Δh associated with the overall process, calculated as the area under the c_p^{exc} curves, depends linearly upon the molar concentration n, confirming that the endothermic peak is mainly related to unfolding, and it is only marginally distorted by aggregation because aggregation is faster at high temperature and affects especially the right part of the thermogram. The aggregation process is typically an exothermic process [11, 19] and, in our conditions, we can estimate that its value has the same order of magnitude of the error on ΔH. A footprint of aggregation is anyway evident in the small shift of the midpoint temperature $T_{1/2}$ upon increasing concentration. The actual $T_{1/2}$ can be estimated by a linear extrapolation of such temperatures to infinite dilution, and yields the

value of $T_{1/2} = 71.1\ ^{\circ}C$ (arrow in Fig. 2.3). The solid line in the inset of Fig. 2.3 gives an estimate of the molar enthalpy associated with the unfolding process: $\Delta H^{exp} = 380 \pm 20\ kJmol^{-1}$. This value is consistent with the previous experiments on lysozyme thermal denaturation [46]. At 50 mg/ml the mean free distance between two molecules is 10 nm, while lysozyme radius of gyration is about 2 nm [6]. Therefore, we may argue that, at low scan rate, unfolded molecules have enough time to building a network of intermolecular bonds, so that molecules cannot refold. On the other hand, at high scan rate, the aggregation remains at a starting phase.

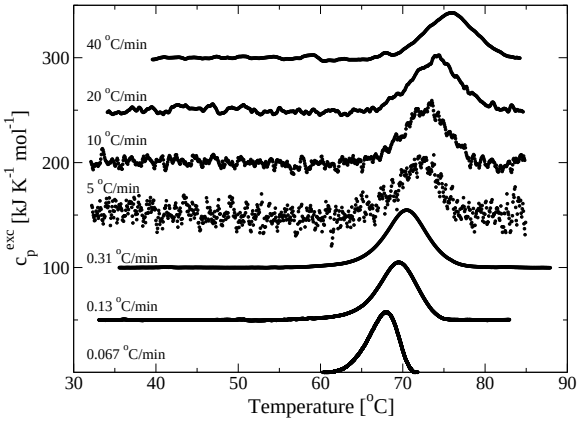

Fig. 2.4: Excess of specific heat (shifted on vertical axis) of neutral buffered lysozyme solutions (54 mg/ml) at different scan rates.

Then, we examined deeper the behavior of only DSC upward scan on 54 mg/ml lysozyme solutions in function of scan rate. We explored a large range of scan rates by using two different instrumental equipments (Fig. 2.4). In both instruments the signal to noise ratio decreased with decreasing scan rate. This effect is more evident in the measurements performed by Perkin Elmer DSC because of small volume used. Besides it was very difficult to subtract the baseline from 0.067 $^{\circ}C/min$ scan up because of, probably, a stronger effect of the aggregation. On the other hand 40 $^{\circ}C/min$ scan is distorted due to a not efficient thermalization. By increasing the rate, the first upward scan shows a drift of the peak towards higher temperatures [197, 220, 208, 157]. We can obtain further informations from these profiles by using appropriate thermodynamic models.

2.2.3 Models for thermal unfolding and analysis of DSC data

Thermodynamic properties of protein solutions have been extensively investigated by differential scanning calorimetry (DSC) techniques. The measured quantity is the *excess heat capacity* (c_p^{exc}) of the solution (see §A.3). Since 1970 the main working hypothesis has been that the unfolding process was reversible [193, 194, 195]. With a few exceptions (ribonuclease A [108] and other particular systems [1, 124]) the protein is not reversible and there are some kinetics effects that characterize thermograms in function of protein concentration, scan rate or other physical parameters [210, 9, 270, 231, 33, 22]. For instance no equilibrium models can explain the scan rate dependence [157]. The main difficulty in the interpretation of DSC data is to evaluate the weight of aggregation and irreversibility in the studied system. There are some criteria to establish that [9, 220]. In some cases the unfolding process is totally irreversible [131, 119, 3], but the main part of physical systems can be studied through mixed models in which reversible and irreversible steps are involved. These cases are generally called *Lumry-Eyring models* [219].

Two-states reversible model

The two-states reversible model is the most simple to describe a conformational change. This model is typically used when thermograms have only a narrow peak because a two state transition is generally fast and cooperative. It can be schematically represented in this way:

$$N \rightleftharpoons U \tag{2.1}$$

where N and U are respectively the native and the unfolded state of the protein.

The most strong assumption is that the two states are always in chemical equilibrium and their relative population is ruled by their difference in the free energy ΔH through the Boltzmann factor.

$$\frac{[U]}{[N]} = K = exp\left(-\frac{\Delta G}{RT}\right) = exp\left\{-\frac{\Delta H_U}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right)\right\} \tag{2.2}$$

where $[U]$, $[N]$, K, ΔG, ΔH_U, T_m and R are respectively the concentration of unfolded and native protein, the equilibrium constant, the difference of free energy and enthalpy in the transition, the temperature at which K=1 and the constant of ideal gas. The apparent excess heat capacity, c_p^{exc}, can

be easily obtained as a function of temperature [219]:

$$c_p^{exc} = \frac{\Delta H_U^2}{RT^2} \frac{K}{(K+1)^2} \tag{2.3}$$

Multiple-states models

As we have already noted, except in rare cases, the unfolding transitions are not purely two-state transitions. Sometimes the experimental thermogram can show a slightly distorted peak. Different hypothesis can be formulated about that and other tests are necessary to try to discriminate the different cases. In fact the deformation could be an effect of irreversibility [219]; in alternative it could suggest the existence of equilibrium intermediates and a deconvolution procedure could help to separate the contributions [81, 163]. Of course, if bumps were evident we could exclude that they are due simply to irreversibility. However bumps could suggest also the unfolding of different protein regions. Typically this interpretation works well in case of big and multi-domain proteins [259]. Sometimes it is really possible to relate an unfolding peak for each domain [112].

Irreversible three-states model

The irreversible model assumes that a protein, after conformational change, cannot restore the correct folding. The reason could be that the unfolded state is much more stable than native or that intermolecular interactions block the folding ("mechanical freezing of an unfolded state"). We considered a more general case with two irreversible states (D_1 and D_2) to consider both effects or consecutive irreversibly unfolded states, more stable than native.

$$N \rightarrow D_1 \rightarrow D_2 \tag{2.4}$$

$$c_p^{exc} = \frac{E_{N \rightarrow D_1}}{v} k_3 K' exp \left[-\frac{1}{v} \int_{T_0}^{T} k_3 K' dT \right] \tag{2.5}$$

$$k_3 = e^{-\frac{E_{D_1 \rightarrow D_2}}{R}(\frac{1}{T} - \frac{1}{T^*})} \tag{2.6}$$

where $E_{D_1 \rightarrow D_2}$ is the activation energy of the transition $D_1 \rightarrow D_2$ and T^* is the temperature at which $k_3 = 1\ min^{-1}$

$$K' = e^{-\frac{E_{N \rightarrow D_1}}{R}(\frac{1}{T} - \frac{1}{T'})} \tag{2.7}$$

where $E_{N \rightarrow D_1}$ is the activation energy of the transition $N \rightarrow D_1$ and T' is the temperature at which $K' = 1\ min^{-1}$ [220].

Lumry-Eyring model

Lumry-Eyring (L-E) models are derived from equilibrium models by adding an irreversible step. They were introduced to correct the distortions of thermogram due to the irreversibility of the unfolding-aggregation process [219]. The most simple L-E model is :

$$N \rightleftharpoons U \rightarrow A \tag{2.8}$$

where A is the irreversibly unfolded state of the protein.

This model supposes that native and reversibly unfolded proteins are in thermodynamic equilibrium, but a fraction of unfolded population is converted into irreversibly unfolded species. The apparent excess heat capacity, c_p^{exc}, is given by

$$c_p^{exc} = \frac{K\Delta H_U}{(K+1)^2}\left(\frac{k_3}{v} + \frac{\Delta H_U}{RT^2}\right) exp\left[-\frac{1}{v}\int_{T_0}^{T}\frac{k_3 K}{K+1}dT\right] \tag{2.9}$$

This expression is more general than the others introduced and can describe them for particular values of parameters.

2.2.4 Analysis of DSC data

In the most complex models the quite high number of parameters makes difficult their correct estimation. Some of them can be correlated and physical information could be not always evident. Thus, tipically the best procedures to reduce these effects are: make simultaneous fits of DSC thermograms collected in function of one parameter [148] or impose some constraints by using physical considerations or obtained in an independent way. We fitted

	Lumry-Eyring				Reversible		Irreversible		Experimental	
Scan rate	ΔH_U	E_{att}	T_m	T^*	ΔH_{rev}	T_m	$\Delta H_{N\rightarrow D_1}$	$E_{D_1\rightarrow D_2}$	ΔH	T_{peak}
0.067	265	371	75.7	71.0	340.9	129.4	2152	-1569	252	68.2
0.13	338	170	73.5	79.4	342.4	159.7	2768	-2348	335	69.6
0.31	376	104	73.0	85.5	343.5	168.2	3020	-2640	370	70.6
5	328	150	77.2	58.1	345.1	144.9	2604	-2179	330	72.6
10	360	127	76.0	56.6	345.9	165.1	2911	-2520	353	73.3
20	373	103	76.2	52.9	347.2	157.7	2936	-2568	343	74.6
40	331	113	79.0	48.6	348.9	125.1	2654	-2300	303	76.5

Tab. 2.1: Comparison among characteristic parameters obtained by experimental thermograms or fitting them through different models.

the experimental thermograms by the last models introduced to compare the goodness of fits and understand which model is the best to describe the thermodynamics of our system. Fig. 2.5 highlights that the best fit is obtained

through the L-E model, so a reversible component and an irreversible one are appropriate to describe the thermal behavior. In Tab. 2.1 the parameters extracted by fits through L-E model show that T_m is quite constant and $< T_m >= 75.2 \pm 1.8\ ^\circ C$. ΔH_U and E_a exhibit a large variability and

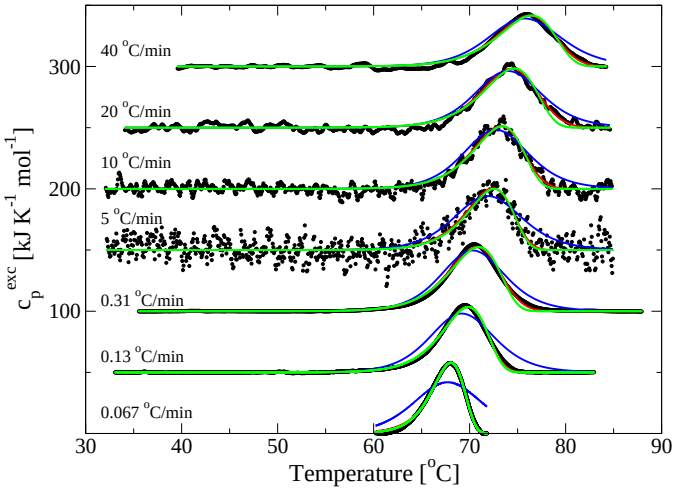

Fig. 2.5: Excess of specific heat of neutral buffered lysozyme solutions (54 mg/ml) at different scan rates (black points). Data are fit by a three-states L-E model (red), by a reversible model (blue) and a three-states irreversible model (green).

are slightly correlated. However, since a particular behavior is not expected as function of the scan rate, we can calculate average values, which are respectively $355 \pm 21\ kJmol^{-1}$ and $131 \pm 29\ kJmol^{-1}$. T^* is subjected to large variations and it is related to ΔH_U and E_a; T^* represents the characteristic temperature of the irreversible unfolding process. All average values are calculated by excluding the values of the extreme scan rates. It is very interesting to highlight that $< \Delta H_{exp} >$ is the same of $< \Delta H_U >$ within experimental errors and has no significative discrepancy with value obtained from study in function of concentration.

2.2.5 Heat induced aggregation of lysozyme

Turbidity experiments

We have carried out turbidity kinetics (Fig. 2.6) upon 18.7 mg/ml lysozyme samples in 0.1 M phosphate buffer at pH 7 and at various temperatures to study more deeply the aggregation process. These kinetics have been made in the temperature range between 60 and 80 $^\circ C$ the relevant range highlighted

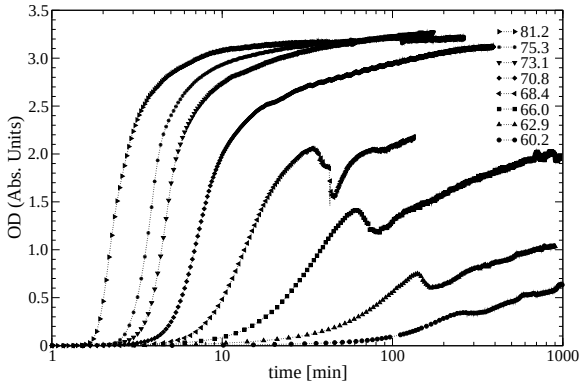

Fig. 2.6: Optical density kinetics on 18.7 mg/ml lysozyme solutions in 0.1 M phosphate buffer at pH 7 within range 60–80 $°C$.

by calorimetric measurements. Over 73 $°C$ kinetics reach a plateau within the thermalization time. At the end of the kinetics these samples appear strongly opalescent and quite hard. Below 70 $°C$ the kinetics are much slower.

Activation energies of folding and aggregation: a clear-cut relation

It is well accepted that there is a strict connection between partial unfolding and aggregation [56], but typically it is considered only in a qualitative way. We tried to understand if in our case we can say something more.

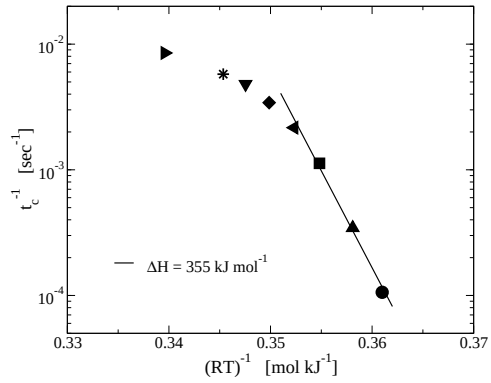

Fig. 2.7: Arrhenius plot obtained from turbidity kinetics in Fig. 2.6: we chose 0.2 absorption threshold to take the characteristic times (t_c) from these kinetics. Black solid line represents $\Delta H = 355\ kJ/mol$.

We obtained characteristic times from kinetics in Fig. 2.6. At these times each kinetic reaches absorption value 0.2. We chose this threshold to be sure that the lower temperature kinetics achieves this value as more possible after thermalization and before precipitation. Fig. 2.7 shows the characteristic times of kinetics (reported in Fig. 2.6) as a function of the incubation temperature. The highest temperatures kinetics have characteristic times lower than thermalization one or a little bigger: that leads to the curvature of the last points and the physical information is misleading. In Fig. 2.7 the slope of first points is related to the enthalpy variation of overall unfolding-aggregation process. Black solid line represents ΔH_U value obtained from fitting procedure of DSC thermograms and it is related to only unfolding process. It is worth to note that black solid line fails to describe points related to the highest temperatures kinetics and this result is not critically due to the particular threshold chosen. The correspondence of the values found in these two different ways suggests that unfolding is necessary to the aggregation and the unfolding rate equals aggregation rate: unfolding is the rate limiting step of aggregation.

Scattering experiments

We prepared a fresh solution of lysozyme 1 mg/ml in 0.1 M phosphate buffer at pH 7, as described in §A.1. We incubated the sample at 61.8 $°C$ and collected the correlation functions as a function of the time. We analyzed the intensity correlation functions by three species (monomers, aggregates and big aggregates). We modelled monomers and big aggregates as simple exponential functions and we took into account a polidispersity [82] for the aggregates (Eq. 2.10).

$$g_2(t) = 1+\beta \left[A_{mon} e^{-\frac{t}{\tau_{mon}}} + A_{agg} e^{-\frac{t}{\tau_{agg}}} \left(1 + \frac{\mu_2}{2!} t^2\right) + A_{AGG} e^{-\frac{t}{\tau_{AGG}}} \right]^2 \quad (2.10)$$

The hydrodynamic radius of monomers $R_{h,mon}$ was fixed at 1.8 nm, but is not critical for the analysis. Hydrodynamic radius of aggregates $R_{h,agg}$ increases quite linearly in time and the increase of the Rayleigh ratio (proportional to scattered intensity) is mainly due to aggregates (Fig. 2.8). At the onset, this measurement is very noisy. Until 75 minutes, the size of aggregates is ill defined because probably their number is small and sample is quite heterogeneous. After 100 minutes the size of aggregates grows and the Rayleigh ratio related to them increases linearly.

Does Rayleigh ratio increase only due to the size of aggregates or also for the increase of their number? In order to calculate the number concentration of aggregates, we need to make an hypothesis on their form factor $P(q)$.

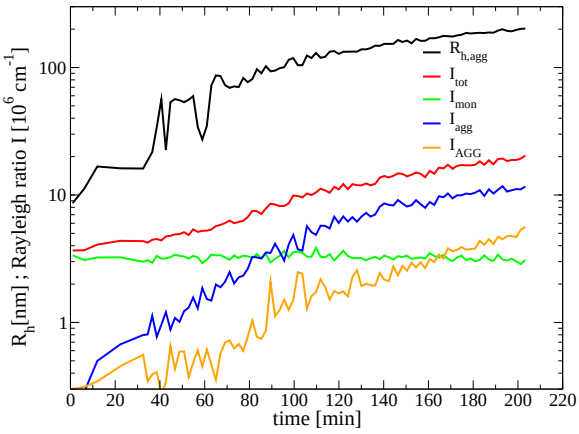

Fig. 2.8: Analysis of DLS correlation functions upon 1 mg/ml lysozyme solution in 0.1 M phosphate buffer at pH 7 at 61.8 $°C$. Temporal behavior of total Rayleigh ratio (red) and fraction related to monomers (green), aggregates (blue) and "big aggregates" (orange). Black points represent the hydrodynamic radius $R_{h,agg}$ of the aggregates.

First, we approximate the gyration radius of the aggregate with its own hydrodynamic radius ($R_{g,agg} \approx R_{h,agg}$). If the aggregate size were very smaller than laser wavelength we would be in the Guinier regime and we could use a fractal dimension $d = 2$ and represent the form factor as in Eq. 2.11.

$$P(R_g, q) = \left[1 + \frac{1}{3}(R_g q)^2\right]^{-1} \tag{2.11}$$

where R_g is the gyration radius of the object and q is the module of scattering vector (Eq. A.6). However, if the size is of the order of magnitude of wavelength, the expression which introduces less artifacts, a minimal model, is the *Fisher-Burford* form factor (Eq. 2.12)

$$P(R_g, q) = \left[1 + \frac{2}{3d}(R_g q)^2\right]^{-d/2} \tag{2.12}$$

It is a generalization of Guinier form factor (Eq. 2.11) because extends a bit the size range covered and can also be used for fractal aggregates with a different compactness. d is called Fisher-Burford fractal dimension. $d = 2$, in addition to make equal Eq. 2.12 and Eq. 2.11, is also used to describe random polymers. However, typically, protein aggregates have a fractal dimension 2.3 and we assume that it is valid also for the aggregates which form in our conditions. By assuming that the aggregates are fractal objects with an

average density equal to the monomer density, we estimate the mean mass of a typical aggregate $\bar{M}_{agg}(t)$ by a scaling equation (Eq. 2.13) [4]:

$$\bar{M}_{agg}(t) = M_{mon}\left[\frac{R_{h,agg}(t)}{R_{h,mon}}\right]^{d} \tag{2.13}$$

where M_{mon}, $R_{h,mon}$, $R_{h,agg}$ are respectively the molecular mass of the monomers, the hydrodynamic radius of the monomers and aggregates.

By using Eq. 2.12–2.13 we calculate the average aggregate mass concentration ($\bar{c}_{agg}$) and average aggregate number density ($\bar{N}_{agg}$) in function of the time:

$$\bar{c}_{agg}(t) = \frac{cM_{mon}[1 - A_{mon}(t)]R(t)}{R_{mon}(0)\bar{M}_{agg}(t)P_{agg}(q,t)} \tag{2.14}$$

$$\bar{N}_{agg}(t) = \frac{cM_{mon}[1 - A_{mon}(t)]R(t)}{R_{mon}(0)\bar{M}^{2}_{agg}(t)P_{agg}(q,t)} \tag{2.15}$$

Fig. 2.9 shows the behavior of $\bar{c}_{agg}(t)$ and $\bar{N}_{agg}(t)$. This qualitative behavior is the same if we assume another fractal dimension (in the range 2.0–2.5), even if the absolute values are different. At the onset of kinetics the mass concentration and the number of aggregates are noisy. The mass concentration increases but the number concentration decreases. We can hypothesize that this behavior is due to the coagulation of big aggregates already formed. This decrease of the aggregates' number is compatible with an exponential

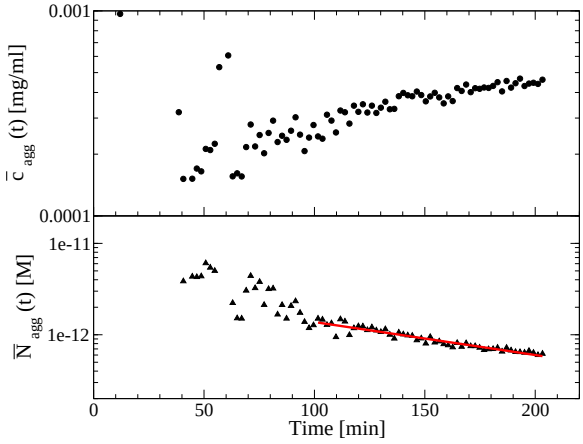

Fig. 2.9: Average aggregate mass concentration $\bar{c}_{agg}$ and average aggregate number density $\bar{N}_{agg}$ vs time for buffered sample. The solid line is an exponential fit (coagulation-like behaviour). Calculations developed by using $d = 2.3$

function, but we cannot be sure about the functional form because of its narrow temporal range.

2.2.6 The role of electrostatic interactions

We evaluated the role of electrostatic interaction by modulating the efficiency of charge screening. We repeated experiments of last section in water solution, corrected at pH 7 by adding small amounts of 1 M hydrochloride acid and without added salt. There is a huge literature on the role of salts in the stability of proteins. We did not consider deeply the specific effects of salts (Hofmeister series [15]), but evaluated their role on solution stability and on protein conformation stability. The addition of salt increases the number of ionic species in solution. This can change a bit the dielectric constant [59], but the main effect is to affect the range of electrostatic interaction. In fact in a salt free solution the electrostatic repulsion is an inverse second power law as a function of intermolecular distance and it is a long range interaction. But the presence of salt increases the ionic strength and makes shorter this range because of a screening effect of charged double layers. Thus, the coulombian interaction is corrected by an exponential damping factor (*Debye-Hückel model*). At high electrolytic concentration (>0.5 M) they can stabilize the proteins by increasing the surface tension and the hydrophobicity of the solvent and stabilizing the core of the protein. However ionic strength can also affect conformation stability in a protein-specific way [41, 40, 152, 59]. An interesting example is a quite huge class of proteins, thermal shock proteins. They are very resistent against cellular stresses (extremes temperatures and ionic strength) and are present in some types of bacteria. For example proteins in *psychrofiles*, *mesophiles*, *thermophiles* and *hyperthermophiles* bacteria can respectively resist to higher and higher temperatures by performing their function [242]. They have respectively higher and higher denaturation temperature. One main structural characteristic which makes different these classes of proteins is the number and the strength of salt bridges on the surface of protein [243]. In fact they are numerous in thermophilic and hyperthermophilic proteins and stabilize the macromolecule by opposing to the unfolding. The increase of ionic strength in solution decreases also the strength of these salt bridges and the propension to unfolding is enhanced [59]. On the other hand, the mesophilic proteins can be stabilized by salts because of screening of close repulsive charges on the protein surface [221, 59]. Salt and other cosolutes or cosolvents can have a *chaotropic* or *cosmotropic* effect. The first creates disorder in the network of water-water H-bonds. The other has the opposite effect. The propension to unfolding is also ruled by balance between intramolecular H-bonds and protein-solvent H-bonds. If the second are dominant the unfolding can be assisted by solvent, also at low temperature. This is the action mechanism of some denaturants or some alcohols.

We have already explained which are the main interactions that contribute to intermolecular potential: *long range* electrostatic repulsions and the *short range* attractions. In particular one aim of ours is to understand the macroscopical effect of competition of interactions. We can guess that in some conditions this equilibrium leads to the formation of finite size equilibrium clusters of proteins (*cluster phase*) [169, 34]. Indeed short range attractions drive proteins closer each other, while long range repulsions avoid the collapse of proteins in amorphous aggregates.

So a key question in our project is to have new insights on the microscopical mechanism of fibril formation. As we referred, the approach we followed was to change the intensity of electrostatic forces (by changing pH: see Chapter 4) and the range of interaction (by changing ionic strength). First we evaluated the effect of change of ionic strength at physiological pH.

DSC experiments at different ionic strength

We want to elicit the role of ionic strength in the unfolding and in the aggregation processes. We already highlighted a strong interconnection between these two processes. One way to modulate the aggregation effects is to change the concentration, but the range of variability is quite limited by instrumental sensitivity, so we cannot use concentrations lower than 10 mg/ml.

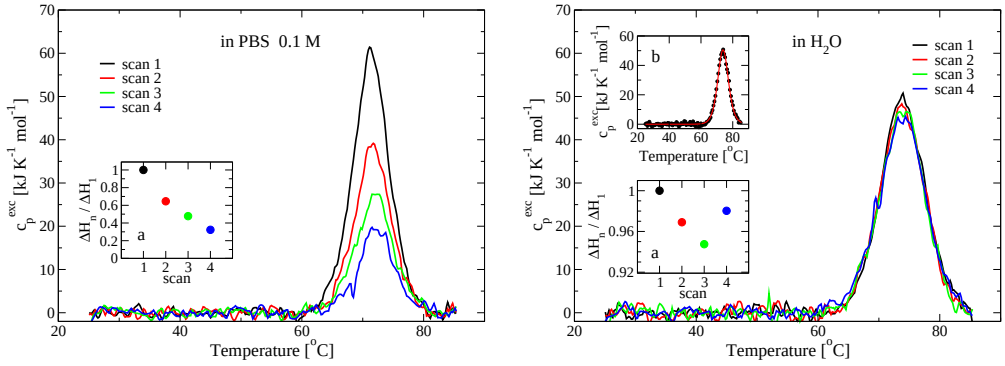

Fig. 2.10: Consecutive DSC scans on lysozyme sample (54 mg/ml) at 20 $°C/min$. Comparison between protein dissolved in phosphate buffer 0.1 M at pH 7 (left) and protein dissolved in Millipore Super-Q water corrected at pH 7 (right). Inset a (in both plots): relative change of ΔH with respect to the first scan. Inset b: fit of one upward scan by a reversible model (Eq. 2.3).

Another way to reduce the aggregation is to decrease the ionic strength because in this case repulsive interactions have a longer range and tend to stabilize the solution. Fig. 2.10 shows the comparison between consecutive DSC scans of two lysozyme solutions at 20 $°C/min$: one prepared in 0.1 M

PBS (left) and one prepared in water corrected at pH 7 by adding small amounts of HCl (right). First, we note that consecutive scans in water give almost the same calorimetric profile, while in buffered solution the thermograms are progressively reduced: insets *a* show the change of enthalpy as a function of number of scans. Further, the inset *b* shows that each scan, in water sample, can be fitted by a reversible model. So this confirms that irreversibility is due to aggregation and that self-assembly is enhanced by increasing the ionic strength. Also the peak temperature tends to increase after each scan, in buffered sample. This is reasonable because the protein is more and more unfolded, so it is necessary to reach an higher temperature, after each scan, to continue the unfolding process. Finally, if we compare ΔH_U in both samples (values obtained through L-E model in PBS), as obtained by fits, we find that ΔH_U is a bit higher in the water sample ($439 \pm 10\ kJ/mol$ in water; $355 \pm 21\ kJ/mol$ in PBS). On the other hand, T_m has no significative discrepancy in both samples ($73.9 \pm 0.1\ ^\circ C$ in water; $75.2 \pm 1.8\ ^\circ C$ in PBS), but the peak temperature of water sample is a bit higher than in PBS. These observations are in accord with the salt destabilizes lysozyme by weakening the salt bridges on the protein surface [234, 59, 253, 243].

Turbidity and scattering experiments at different ionic strength

We carried out turbidity kinetics (Fig. 2.11) of 18.7 mg/ml samples without salt and in the range 60–80 $^\circ C$ to understand the role of ionic strength in the aggregation process.

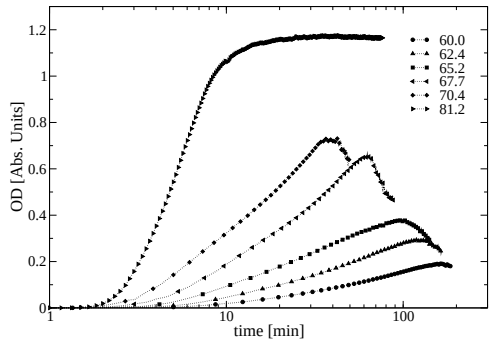

Fig. 2.11: Optical density kinetics on 18.6 mg/ml lysozyme unbuffered solutions at pH 7 and at different temperatures within the range 60–80 $^\circ C$.

Over 70 $^\circ C$ kinetics reach the plateau essentially within the thermalization time. Samples appear liquid and turbid so the plateau represents only a temporary equilibrium between aggregation and precipitation. Below 70 $^\circ C$

kinetics are slower and eventually, the samples appear a little opalescent and show aggregates on the bottom of the cell. These kinetics reach absorption values less than in salt solutions. Visual observations allow us to say that aggregates are smaller in the unbuffered solution, but enough big to precipitate. Growth rates are also slower because electrostatic repulsions are not screened by salt and tend to keep far the molecules each other. Likely, the entropy provides for the clustering and the hydrophobic effect which binds the proteins. A similar DLS measurement and analysis, already performed for buffered lysozyme solution, has been led for the free salt sample too. The

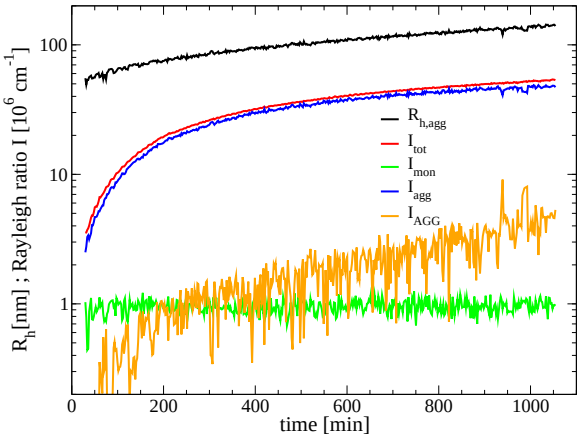

Fig. 2.12: Analysis of DLS correlation functions of 1 mg/ml lysozyme in water at pH 7 at 61.8 $°C$. Temporal behavior of total Rayleigh ratio (red) and fraction related to monomers (green), aggregates (blue) and "big aggregates" (orange). Black points represent the hydrodynamic radius of the aggregates.

hydrodynamic radius of the monomers was fixed at 1.8 nm but it is not critical for the analysis. The hydrodynamic radius of aggregates is around 50 nm after 30 minutes of incubation and increases (Fig. 2.12). This confirms our conclusions about turbidity kinetics and visual observations. The growth of Rayleigh ratio is due almost totally to the aggregates. In the unbuffered sample the hydrodynamic radius of aggregates increases with less noise. At low ionic strenght, the coagulation starts after than in the buffered solution and the coagulation rate is slower. The repulsive interactions are stronger so the strong aggregation is inhibited. In particular, elongated structures or more branched (Fig. 2.14) are preferred to maximize the distances between charged compact aggregates. Fig. 2.13 shows the behavior of $\bar{c}_{agg}(t)$ and $\bar{N}_{agg}(t)$. We can note a similar behavior than Fig. 2.9. In Fig. 2.13 we can say something more about the coagulation because, in water, this process

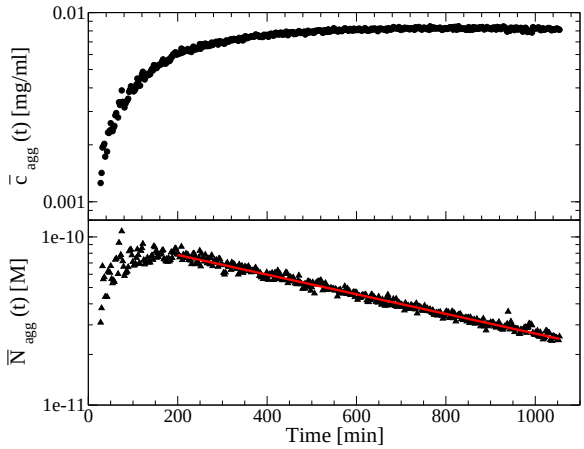

Fig. 2.13: Average aggregate mass concentration $\bar{c}_{agg}(t)$ and average aggregate number density $\bar{N}_{agg}(t)$ vs time in water samples. The solid line is an exponential fit (coagulation-like behaviour). Calculations developed by using $d = 2.3$.

takes more time before multiple scattering occurs. In particular the number density decreases as an exponential function.

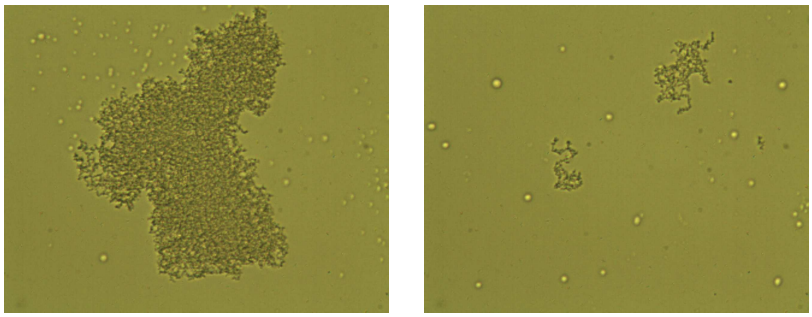

Fig. 2.14: Optical microscopy images (40X) of aggregates from a drop of 1 mg/ml lysozyme free salt solution (pH 7), after incubation at 70 $°C$.

Fig. 2.14 shows the typical sizes aggregates of the unbuffered sample (obtained by an inverted microscope with colour corrected infinity optical system *MOTIC AE21*, equipped with 1.3 MP digital camera).

We also built the Arrhenius plot (Fig. 2.15) by turbidity measurements in water (Fig. 2.11), just to compare the obtained result to measurements in buffer. Because of low turbidity signal, choosing a threshold enough far from plateau values is difficult, but we chose 0.1. The curvature of Arrhenius plot starts very early. Only first two-three points seem reliable, but the slope between them is compatible with DSC measurements ($\Delta H_U = 439\ kJ/mol$).

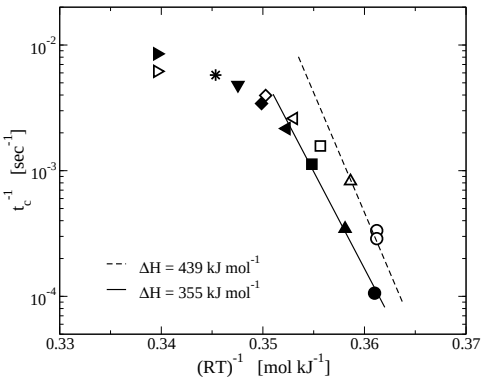

Fig. 2.15: Comparison between Arrhenius plots obtained from turbidity kinetics with buffered solutions (0.1 M) and without salt. The full points are the same of Fig. 2.7. We chose absorption threshold 0.1 for kinetics without salt in Fig. 2.11 (empty symbols).

2.2.7 Conclusive remarks: electrostatic vs hydrophobic interactions

Our aim is to relate the different aggregation pathways with a different balance among the main intermolecular interactions of proteins. In this section we decided to enhance the hydrophobic contribution with respect to the electrostatic one, by increasing the temperature. In this way the proteins partially unfold and their aggregation is favoured. We studied the thermodynamic properties of lysozyme at pH 7 and showed the irreversibility of native-unfolded transition (Fig. 2.2-2.5). The shift of the peak temperature as a function of the scan rate highlighted that the unfolding transition is affected by aggregation because they are temporally coupled processes (Fig. 2.4). On the other hand, aggregation is ruled by unfolding: in particular, we showed experimentally and quantitatively that unfolding is the rate limiting step of aggregation (Fig. 2.7). Besides we evaluated the role of ionic strength on the unfolding-aggregation process. DSC experiments demonstrated that irreversibility is due to aggregation (Fig. 2.10). We highlighted that the decrease of ionic strength (from 100 mM to $<$10 mM) changes a bit the main thermodynamic parameters (T_m and ΔH_U) suggesting, on a microscopical scale, the stabilization of the protein. This can be explained as the strengthening of surface salt bridges, because of reduction of salt screening [243] (Fig. 2.10). Thorough calorimetric study, we focused on the aggregation process around 60 $°C$, where lysozyme is slightly unfolded. In this

way the process is slow enough to be studied on very small scales. First we elicited that aggregated structures are similar in buffer and in water (compare Fig. 2.14 with Fig. 2.17), but at low ionic strength the aggregates are smaller and the kinetics is slower (Fig. 2.8-2.12). DLS kinetics also reveal the rapid formation of aggregates in few tens of minutes. In particular, at a given time, these aggregates are bigger in the buffer solution than in water, so the role of ionic strength, on a mesoscopic scale, is to accelerate the rate of formation of the clusters. Finally, we found that the coagulation between objects formed at the onset of kinetics is favoured with respect to monomer addition (Fig. 2.9-2.13). From a qualitative point of view, this is easily understood. If we compare the electrostatic and the hydrophobic contributions, one obtains that this balance can change naturally by increasing the aggregate's size. In fact charge of aggregates cannot increase indefinitely by their size and the surface charge density reaches a saturation value, as a first approximation [12]. However the entropic gain, to bury the hydrophobic regions to the solvent, is proportional to faced surfaces. This contribution depends upon size and pushes the aggregates to stick, preferentially, with other aggregates rather than monomers [42]. At these temperature the kinetics information is distorted by the coupling between the kinetics of unfolding and aggregation. In order to delete this trouble, we will study the aggregation of unfolded proteins in the next section.

2.3 Irreversible gelation of thermally unfolded lysozyme

2.3.1 Introduction: aggregation of an unfolded protein

As we concluded in § 2.2.5 the increase of aggregation rate with the incubation temperature is strongly related with the increase of the unfolding rate. The higher is temperature, faster is the transition native-unfolded and intermediate states have a shorter life. Around denaturation temperature, from a time point of view, proteins go through different intermediate states before arriving in the unfolded state, eventually aggregated. This is equivalent, from an ensemble point of view, to the equilibrium among different intermediates with time dependent equilibrium constants [218]. In these conditions the unfolding kinetics and the aggregation kinetics are coupled and can affect each other. Thus, we decided to study more deeply lysozyme solutions at pH 7, only at high temperature (80 $^{\circ}C$), where calorimetric data show no more thermodynamic processes. In fact, at high temperature the conversion from native to unfolded state is more rapid and we can guess that the kinetics we

observe characterize mainly the self-assembly process of "unfolded" proteins. Our first conclusion was that thermodynamic irreversibility, elicited in the second calorimetric upward scan, is due to aggregation (see Fig. 2.2). Thus the unfolded proteins are stuck into supramolecular aggregates and cannot recover their native structure. Also, some works suggest that in our conditions the aggregation of lysozyme is due to hydrophobic effect enhanced by unfolding. In fact the protein hydrophobic surface exposed to the solvent changes rapidly between 75 and 85 $°C$ [123, 255]. Instead the secondary structures in the same temperature range are altered only upon long incubation [173]. The breaking and the exchange of the disulfide bridges cannot have any roles in this process because a reducing agent is required [107, 269].

2.3.2 The aggregation process

We prepared fresh solutions of lysozyme in 0.1 M phosphate buffer at pH 7 and studied the aggregation process through turbidity measurements (Jasco UV-VIS spectrophotometer) in function of protein concentration in the range 0.2–20 mg/ml at 81 $°C$. The optical density reaches a plateau, except the lowest concentration samples, affected by precipitation [204].

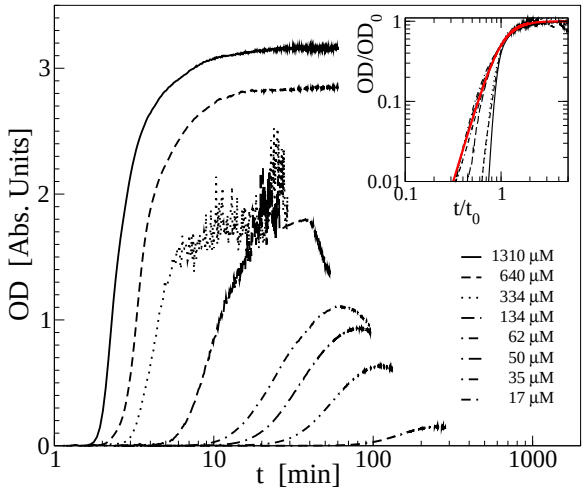

Fig. 2.16: Optical density kinetics of lysozyme solutions at different protein concentrations and incubated at 81 $°C$. Inset: scaling of kinetics by OD_0 and t_0. The bold red curve is a fit to Eq. 2.16

The curves in Fig. 2.16 exhibit a similar shape, if one excludes the early kinetics at the highest protein concentrations, which lies within the thermalization time. Such a similarity indicates that the same mechanism operates

at all the concentrations. Indeed, the OD curves may be rescaled by identifying a characteristic time t_0 and a plateau value OD_0, as shown in the inset. The bold red curve in the inset of Fig. 2.16 is a sigmoidal Hill curve:

$$\frac{OD_0}{OD} = 1 + \left(\frac{t_0}{t}\right)^n \tag{2.16}$$

The latter expression with an exponent $n = 4$ reproduces approximatively the growth of the OD signal at all the concentrations. At very high concentration the initial part of the kinetics is not reliable since the growth is almost completed during the sample thermalization. In the present experiments the final plateau is not barely due to monomer or aggregates consumption [75]. Other effects, such as multiple scattering or mesoscopic gelation, may affect the late stages [21].

Aggregate formation: turbidity and characteristic times

We plotted suitable characteristic times t [204] of turbidity kinetics and we found that they exhibit an inverse quasi-linear dependence upon molar concentration c. More precisely, the data can be fit with a power law $t \approx c^{-1.15}$ [175]. A linear dependence beetween the aggregation rate and the protein concentration calls for a second order reaction kinetics, where the growth is determined by the coalescence of two objects (two monomers, two aggregates, one monomer and one aggregates, an so on) [173]. This suggests that this aggregation process is not specific and there are not particular intermediates which trigger the aggregation [173].

Aggregate structure: optical microscopy, AFM and small angle light scattering

We observed the morphology of aggregates after incubation at 81 $°C$ by optical microscopy. We observed directly the samples on a glass sheet without further procedures. The images in Fig. 2.17 show large clusters with a size spanning from a few tens to hundreds of micrometers, made of smaller micron-sized units [204]. The heterogeneity of shapes and sizes is a clear result of a caotic and not specific coagulation process.

In order to get more quantitative, and statistically significant, informations about the structural properties of such aggregates, we performed small angle light scattering (SALS) measurements on incubated samples at different concentrations (Fig. 2.18). In each circumstance, the incubation time is taken as the time when the OD signal reaches a plateau value. At the highest concentration in Fig. 2.18, we clearly observe a power law dependence in the

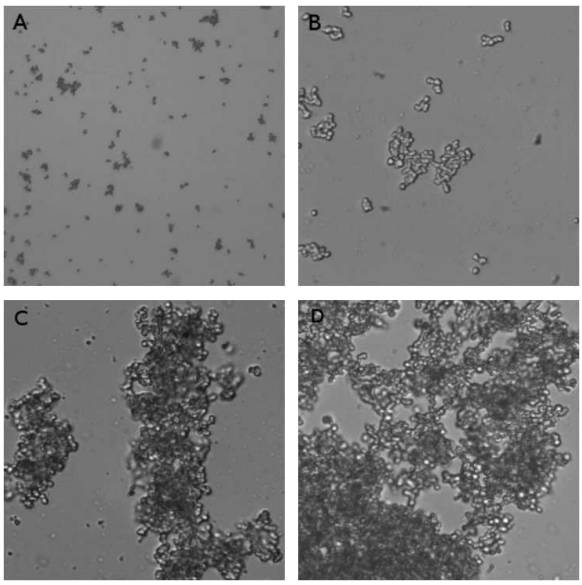

Fig. 2.17: Optical microscopy images of a lysozyme solution at different concentrations after incubation at 81 $°C$. a) $c = 50\ \mu M$, 336.5X336.5 μm^2 (10X); b) $c = 50\ \mu M$, 91.2X91.2 μm^2 (40X); c) $c = 134\ \mu M$, 91.2X91.2 μm^2 (40X); d) $c = 334\ \mu M$, 91.2X91.2 μm^2 (40X).

range of scattering vector between 0.1 and 1 μm^{-1}. The curve is bent at higher scattering vectors. Such a profile of the structure function is typical of clusters having a radius of gyration R_g of tens of micrometers (consistent with microscopy images), which are made from smaller units with a size R of a few microns. The shape of the structure functions in the high range of scattering vectors q is substantially conserved at all the concentrations, while the power law region is restricted to a limited range by lowering the protein concentration, showing that the observed size of clusters R_g is lower at lower concentrations.

Guided by the microscopy images, we fit the scattered intensity data $I(q)$ by using a hierarchical form factor which includes a form factor for the small compact object $F(q)$ and a structure factor for the large cluster $S(q)$ [36]:

$$I(q) = S(q)F(q) \tag{2.17}$$

We used a general function $S(q)$ developed to model fractal aggregates with fractal dimension D and radius of gyration R_g, made from m smaller compact

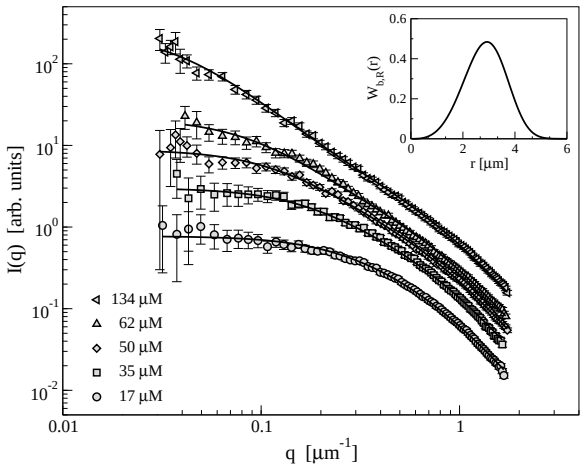

Fig. 2.18: Small Angle scattered light from lysozyme solutions after incubation at 81 $^{\circ}C$ at different concentrations. The samples are diluted to obtain a transmittance of about 90%. The lines are the best fits and the inset shows the used radius distribution of the small aggregates.

objects of size R [170]:

$$S(q; R_g, R, D) = \frac{1}{m}\left[1 + \frac{m-1}{\left[1+(\xi q)^2\right]^{\frac{D-1}{2}}}\frac{\sin\left((D-1)\,\mathrm{atan}\,(\xi q)\right)}{(D-1)\,\xi q}\right] \qquad (2.18)$$

where $\xi = R_g/(D(D+1)/2)^{0.5}$ is the correlation length related to a cluster, and m is the number of small units which are assembled in one cluster: $m = 1 + \Gamma(D+1)\,(\xi/R)^D$, where $\Gamma(x)$ is the Gamma function.

The latter expression introduces a very minimalistic model for the structure of aggregates. Indeed, we obtain only three parameters: the size of clusters R_g, the size of the constituent units R, and the fractal dimension D, which is essentially a measure of the packing of the small units into the large aggregates. In the present case, the value of $D = 2.0$ has been fixed to the best fit value of the curve at the highest concentration. The best fit values for the radius of gyration R_g of large clusters are 6.2, 9.3, 13, 18 and 45 μm respectively for the concentrations $c = 17$, 35, 50, 62 and 134 μM, consistent with microscopic observations.

The form factor of the small units $F(R, q)$ may be fit by using either the well known expression for hard spheres or the Fisher-Burford expression $P(r, q)$ (Eq. 2.12) [170]. We fit the data of Fig. 2.18 by allowing a polydispersity in the radius R of the smaller units, and by using a Weibull distribution

function $W_{b,R}(r)$ with average R [36, 179].

$$F(R,q) = \int_0^\infty W_{b,R}(r)P(r,q)\mathrm{d}r \tag{2.19}$$

$$W_{b,R}(r) = b\left[\frac{\Gamma(\frac{1}{b}+1)}{R}\right]^b r^{b-1}\exp\left\{-\left[\frac{r\Gamma(\frac{1}{b}+1)}{R}\right]^b\right\}$$

The parameter b is a measure of radii polydispersity. The best fit parameters are $R = 2.85\ \mu m$, $b = 4$, and $d = 3$. The value $d = 3$ marks the presence of very compact aggregates and this is the very cause of thermodynamic irreversibility [204].

By optical microscopy and small angle light scattering, we are addressing the structural properties of lysozyme aggregates on a microscopic length scale. In order to extend the structural details to a smaller length scale, we performed atomic force microscopy experiments on lysozyme aggregates formed at high temperature (Fig. 2.19). In particular, we incubated at 81 $°C$ a 35 μM lysozyme solution in 0.1 M phosphate buffer at pH 7 for two hours (the necessary time to reach the plateau of optical density). Then we put few microliters of sample on a mica plate, we washed the sample dropwise by Millipore SuperQ water and dried with a gentle stream of dry nitrogen. Fig. 2.19

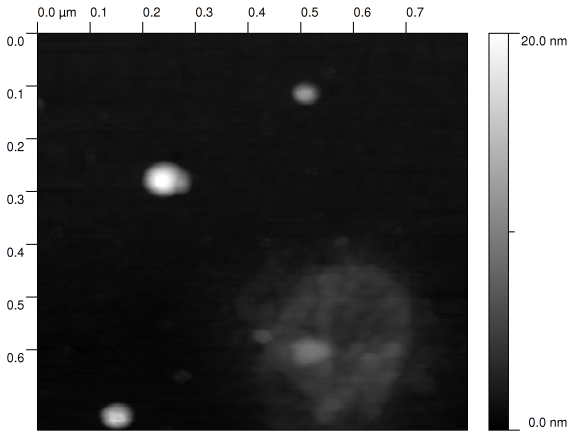

Fig. 2.19: AFM image of a 35 μM lysozyme solution after incubation at 81 $°C$. The side length is 800 nm. The side bar indicates a grey tone mapping of heights.

displays some typical aggregates. They are compact aggregates with a typical size of a few tens of nanometers. Their aspect mirrors the morphology of the aggregates found on larger length scales by optical microscopy.

2.3.3 Mechanical properties of aggregates

Viscoelasticity experiments and analysis

After incubation at 81 $°C$ lysozyme solutions become viscous as expected when large clusters are formed. At concentrations higher than 350 μM the samples exhibit no macroscopic flow, and they appear as turbid macroscopic gels. The viscoelastic properties of such incubated samples were investigated by measuring the viscoelastic spectra at the different concentrations and at 20 $°C$. The *storage* G' and *loss* G'' *moduli* are shown in Fig. 2.20.

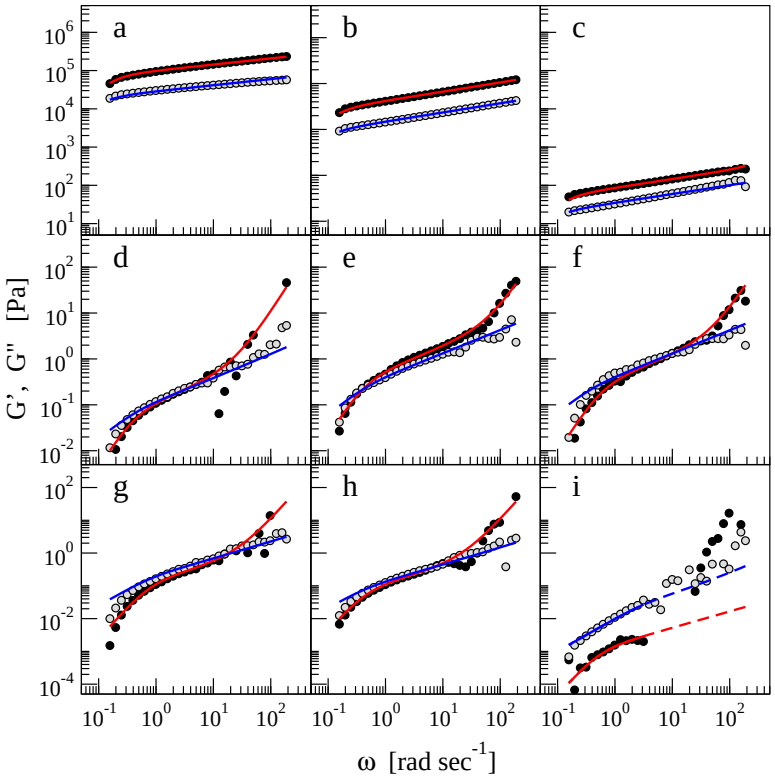

Fig. 2.20: Viscoelastic spectra of lysozyme solutions after incubation at 81 $°C$ at different concentrations: *storage modulus* G' (black circles), *loss modulus* G" (grey circles). The lines are best fits to Eq. 2.20–2.21. a) 4000 μM, b) 1.310 μM, c) 640 μM, d) 334 μM, e) 134 μM, f) 62 μM, g) 50 μM, h) 35 μM, i) 17 μM.

The most evident result is that, at the high concentrations, the viscoelastic spectra are almost flat and they have much higher values than the spectra at the lower concentration. This is essentially a quantitative confirmation of the visual inspection of firm gels at the higher concentrations.

Also in the set of lower concentrations, both the storage and loss spectra display a power law behavior with about the same exponent $n = 0.5$ in the range of strain frequencies ω between 1 and 10 $rad\ sec^{-1}$. We may affirm that in this frequency range the sample exhibits a gel-like behavior [72, 266]. At lower frequencies, the moduli are bent towards lower values, that is they recover the typical liquid behavior and display macroscopic flow. The shape of the viscoelastic moduli are alike to the so called *Rouse* model, which describes the mechanical relaxation of a sample of not entangled clusters [216]. At higher frequencies, the storage modulus G' rises with an asymptotic square dependence upon the frequency. We fit the data by using the following *ad hoc* expressions:

$$G' = G'_0 \frac{(\tau'_0\omega)^2}{1+(\tau'_0\omega)^2}\left[1+(\tau'_0\omega)^2\right]^{\frac{n}{2}} + G_1(\tau_1\omega)^2 \tag{2.20}$$

$$G'' = G''_0 \frac{(\tau''_0\omega)}{1+(\tau''_0\omega)^2}\left[1+(\tau''_0\omega)^2\right]^{\frac{n+1}{2}} + \eta_b\omega \tag{2.21}$$

In the first term of the latter two equations, the fraction is the usual expression for a Maxwell liquid [72], while the expression in square brakets, with the parameter $n = 0.5$ in the exponent, is a suitable term that reproduces the Rouse behavior. In the three highest concentrations the spectra are more flat with $n = 0.20 \pm 0.01$. The parameter G'_0 and G''_0, and τ'_0 and τ''_0 were left free in the fit and they have consistent values. The average values $G_0 = (G'_0 + G''_0)/2$ and $\tau_0 = (\tau'_0 + \tau''_0)/2$ are displayed in Fig. 2.22. The second term in Eq. 2.21 is introduced to include the effect of bulk viscosity η_b, which is of the order of water viscosity, that is $\eta_b = 10^{-3}\ Pa\ s$. In these measurements, the addition of this term is almost irrelevant, since the loss moduli have quite high values. The second term in the Eq. 2.20 for the storage modulus is included to fit the increase of G' at high frequencies. Interestingly, the parameter $G_1\tau_1^2$ is the same at all the concentrations and has a value of $10^{-3}\ Pas^2$. This suggests the existence of a relaxation process at small time scales and hence at small length scales, which occurs at all the concentrations [202]. This observation is also consistent with the results of viscosity experiments.

Viscosity experiments and analysis

We measured viscosity as a function of shear rate on the incubated samples. We found a typical *shear-thinning* behaviour (Fig. 2.21). The curves of Fig. 2.21 were fit by the well known *Carreau* expression [18]. At concentration higher than 350 μM, when the flow is completely inhibited, we added a

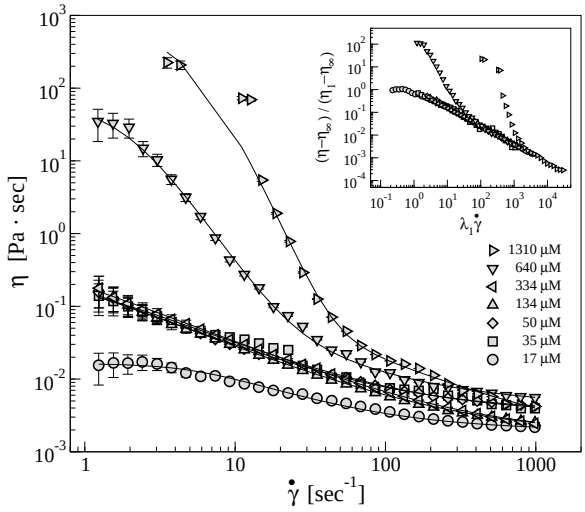

Fig. 2.21: Viscosity vs shear rate of lysozyme solutions after incubation at 81 $^{\circ}C$ at different concentrations. The solid curves are fits to Eq. 2.22. Inset: scaling of viscosity curves according to Eq. 2.22

second term, which extends the Carreau model:

$$\eta = \eta_\infty + (\eta_1 - \eta_\infty)\left[1 + (\lambda_1\dot{\gamma})^2\right]^{\frac{n_1-1}{2}} + \eta_2\left[1 + (\lambda_2\dot{\gamma})^2\right]^{\frac{n_2-1}{2}} \qquad (2.22)$$

Eq. 2.22 reproduces the shear thinning with the exponent $n_1 = 0.2$, and yields two useful parameters: the zero-shear-rate viscosity $\eta_0 = \eta_1 + \eta_2$, and the infinite-shear-rate viscosity η_∞. The parameter λ_1 is a characteristic time. Its value is of the order of a few seconds, consistent with the values of τ measured in the viscoelastic spectra. The curves in Fig. 2.21 have been rescaled by using the fit parameters λ_1, η_1 and η_∞. All the curves collapse into the same master curve, with the exception of the second relaxation process present at the two highest concentrations.

Cluster formation and percolative transition

In Fig. 2.22 we plot the behavior of main rheological parameters in function of protein concentration. G_0 is a representative measure of elastic modulus of viscoelastic spectra and η_0 is the zero-shear-rate viscosity. The viscoelastic spectra have the same dependence by concentration, in fact these two parameters reflect the macroscopic gel behavior of the samples [204]. On the other hand, the infinite-shear-rate viscosity η_∞ does not depend upon protein concentration. This highlights that on a microscopic length scale,

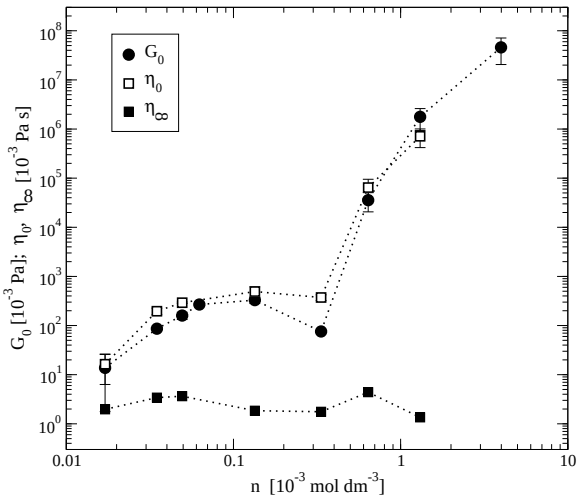

Fig. 2.22: Rheological parameters vs protein concentration: G_0 (full circles), η_0 (empty squares), η_∞ (full squares): see text.

the mechanical properties are similar in the entire range of explored concentrations. The same concentration-independent behavior was found in the structure and size of small micron-sized aggregates that form larger clusters (Fig. 2.18). At high shear rate, the large clusters are destroyed and the small compact irreversible aggregates remain at all the concentrations. At low shear rate, the cluster structure and size depend upon the initial concentration, as found in the SALS structure functions (Fig. 2.18). This suggests that interactions which bound the large clusters are weaker than interactions which bound the small compact aggregates. The elastic modulus and zero shear-rate-viscosity dramatically increase above the threshold of 340 μM lysozyme concentration. This reveals that the formation of a macroscopic gel is due to the percolation of large clusters, which occurs above a given threshold concentration. Our findings confirm and make stronger the ideas suggested from some recent experiments. According to Krebs [129] the formation of particulate gels might be a generic property of all proteins and polypeptidic chains, as already proposed for amyloid fibrils [43].

In the context of our work, we may note that the different aggregation processes (amorphous self-assembly, fibrillation, crystallization) are the macroscopical effect of a particular microscopical balance of opposite interactions. Of course, the particular primary structure of a protein and its tridimensional folding give some restrictions to the solvent conditions that we need to have a certain balance between the involved interactions.

2.3.4 Conclusive remarks

We investigated the aggregation process of thermally denatured lysozyme. In this way we avoid that the unfolding kinetics interferes with the aggregation kinetics. Scaling of turbidity kinetics allows us to conclude that aggregation mechanism is independent upon concentration (Fig. 2.16). SALS measurements revealed a hierarchical structure of the formed aggregates: micron sized and compact aggregates, present at all concentrations, are assembled into suprastructures, which increase their size by increasing the lysozyme concentration (Fig. 2.18). This behavior is confirmed by rheological measurements, which have clearly identified a transition threshold (around 5 mg/ml) beyond which the system forms a gel. This is called *critical gel concentration* and the overall mechanism falls within the framework of an heat induced *sol-gel*transition [92] (Fig. 2.22). Also viscosity measurements highlighted this dependence upon concentration but it is interesting to note that the strong mechanical stress can destroy the suprastructures, but not the smallest aggregates. So it seems that the interactions which bind the aggregates on different scales are different in strength. In this chapter we deeply discussed about the role of temperature on the aggregation. This physical parameter modify the structure of the protein in a not specific way. In the next chapter we will deal with the effect of an alcohol on the stability of lysozyme, on its conformation and on its aggregation pathway.

Chapter 3

Lysozyme stability affected by alcohols

3.1.1 Introduction: the role of osmolites and alcohols in affecting protein conformation and stability

Proteins are more soluble and perform their function when folded in their native conformation. In turn, the native state is the result of several and different types of interactions [244]. Besides proteins are marginally stable, [248] because the stability of a protein and of a solution are ruled by a delicate balance of interactions [32]. The existence and the characterization of intermediates, partially folded protein conformations, are particularly interesting for the study of protein folding and stability [192]. Indeed, the protein folding follows a precise sequence of intermediate states through a progressive decrease of the free energy [55]. These states have a very short lifetime, so it is very difficult to study them. In most cases, only the unfolded and the native state are well populated [205].

Specific cosolutes can alter their stability and can have a very central role in the kinetics of folding-unfolding or formation of particular supramolecular structures. These additives are mainly classified in two groups: *stabilizers* and *denaturants* [103]. Stabilizers include sugars, osmolytes, kosmotropic salts, alcohols; denaturants include guanidine, urea, chaotropic salts.

In particular we are interested to study the effect of alcohols, 2-2-2 trifluoroethanol (TFE) in particular, on lysozyme solutions because alcohols are well-known to affect the protein stability and the intermolecular interactions. Their preferential solvation of the protein limits the configurational space accessible to the solvent [251, 38, 252, 189] and the entropic contribution is lower [50]. Alcohols cause also the decrease of the bulk dielectric constant, so the electrostatic repulsion is enhanced [50, 120, 38] and the solu-

tion is more stable. Thus, alcohol concentration can modulate the interaction contributions in different ways and drive the aggregation pathway in different directions. Alcohols are often used to accelerate the folding rate [145, 44, 132] or stabilize intermediate conformations [223], or act as osmolytes [120]. High concentrations of alcohol can destabilize the protein leading, in some cases, to the formation of fibrils [130, 64, 142]. Reduction of the solvent accessible area can also cause a decrease of the heat capacity difference Δc_p [164] and a correlated decrease of the unfolding temperature [245, 257].

The same effect has been also recently shown to affect protein *cold denaturation* [153]. Another explanation for the reduction of the unfolding temperature has been attributed to an increase of the unfolding entropy due to the preferential binding of alcohols to the protein [177].

In particular, TFE can form intra- and intermolecular hydrogen bonds [206] and operates a preferential solvation of certain backbone groups [265, 211, 38]. For this reason TFE can stabilize α-helix [28, 232, 147, 60] and unfolding intermediates, it can also alter the pathway of protein folding [145, 44, 102, 31, 85] and affect the protein aggregation [38, 141, 97] with a behavior depending upon alcohol concentration [113].

3.1.2 Lysozyme solvational properties

We studied the effects of TFE on the molecular conformation and on the thermal unfolding of hen egg-white lysozyme. The used sample was lysozyme in phosphate buffered solution at pH 7.

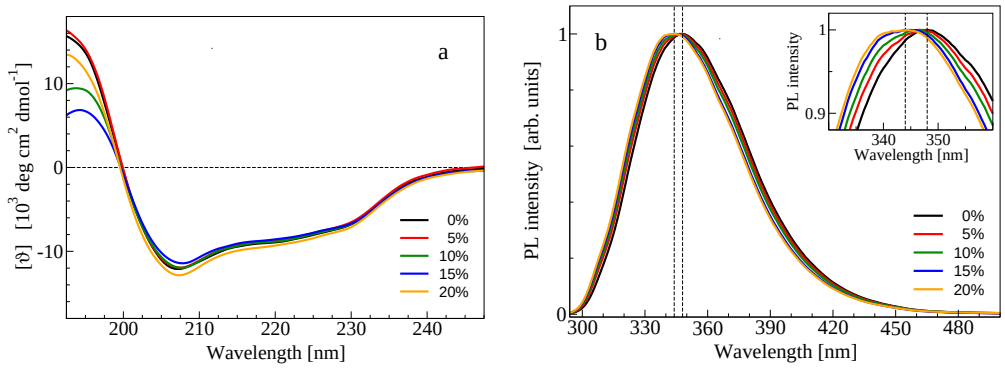

Fig. 3.1: *a*) Far UV ellipticity of lysozyme solutions ($c = 0.2\ mg/ml$) at 20 $°C$ and different TFE concentrations. *b*) Steady state fluorescence spectra of lysozyme solutions ($c = 0.5\ mg/ml$; $\lambda_{ex} = 290\ nm$) and measured at 10 $°C$ for different TFE concentrations. The inset shows a zoom of the emission spectra to highlight the red-shift of the peak.

Far UV CD spectra show that secondary structure is slightly affected by

TFE content (range $0-20\%$) at room temperature, but steady state fluorescence of tryptophans shows a big blue-shift of the emission band (Fig. 3.1). This highlights that Trp62 (this is the only one exposed to the solvent in the native state) feels a not polar environment [133]. Time resolved fluorescence measurements (data not shown) confirm the preferential solvation of Trp62 and show a red-shift in function of time: this curve can be fit by a double exponential at all alcohol concentrations. Thus, we can conclude that TFE, at room temperatures, does not change secondary structure and affects slightly the tertiary structure, because of preferential hydration.

3.1.3 Thermal unfolding

At higher and higher temperatures, tryptophan emitted intensity decreases because of the activation of some not radiative channels from excited electronic state (Fig. 3.2a). Besides the sigmoidal red-shift is due to the exposure of Trp108 to the solvent (Fig. 3.2b). This residue, in the native state, is localized near the active cleft, between α and β domains, so the red-shift suggests a slight tertiary conformational change in this region, because the red-shift is due to exposure to a polar solvent [133]. The blue-shift, already highlighted at low temperature, is also present at high temperature ($90\ ^{\circ}C$), even if the effect is very small. The unfolding temperature is reduced by increasing TFE.

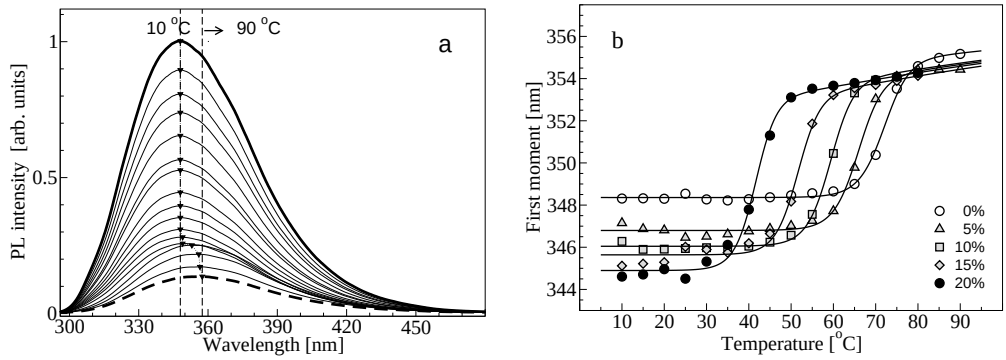

Fig. 3.2: Panel *a*: Steady state fluorescence spectra of lysozyme solutions ($c = 0.5\ mg/ml$; $\lambda_{ex} = 290\ nm$) at different temperatures between $10\ ^{\circ}C$ (bold solid line) and $90\ ^{\circ}C$ (bold dashed line). The black triangles indicate the position of band peaks. Panel *b*: Temperature dependence of the first moment of emission bands as a function of TFE content. Solid lines are the result of fitting procedure throw a three-states model (Eq. 3.1-3.2-3.3).

CD spectra in function of the temperature and TFE concentration elicit an isodichroic point around 204 nm, at lower TFE contents. That sug-

gests a cooperative two-states transition. At higher TFE concentrations, the isodichroic point is not more sharp, expecially at 20 % TFE (Fig. 3.3).

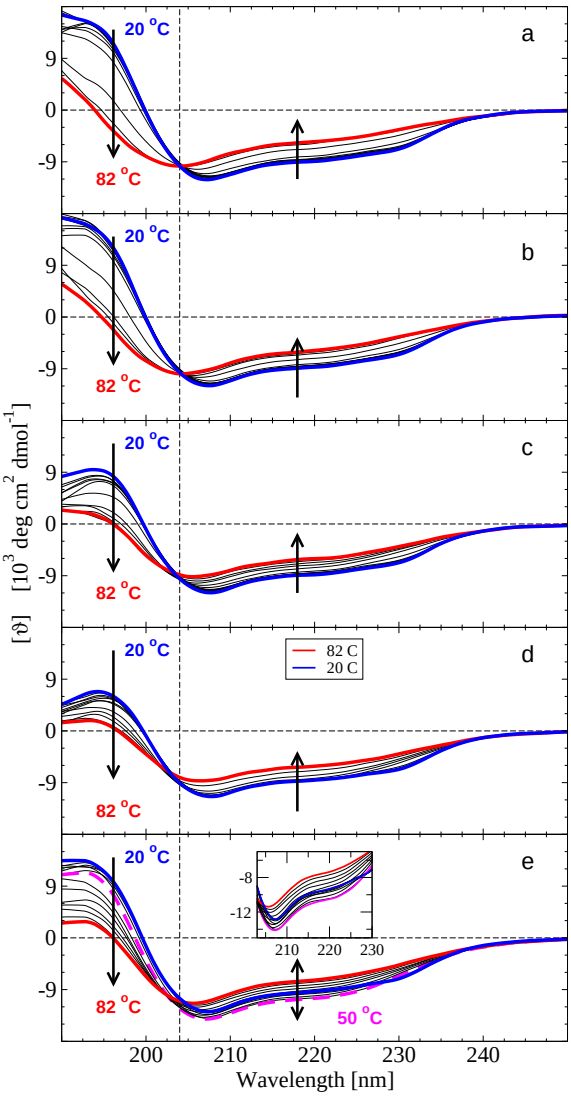

Fig. 3.3: Panels a) b) c) d) e) show the thermal behavior of far UV CD spectra of lysozyme solutions with 0, 5, 10, 15 and 20 % of TFE, respectively. The blue and red solid lines indicate the spectra at the lowest (20 $°C$) and the highest (82 $°C$) temperature. The dashed magenta line in panel e) indicates the spectrum at 50 $°C$. The vertical dashed reference indicates the spectral position at 204 nm. The inset in panel e) is a zoom.

3.1.4 Thermal stability: analysis by a three-states model

During folding a protein goes towards the native state through different states (*kinetic intermediates*). Proteins are marginally stable, so the conversion from native state to some conformations close in the energy landscape (not the conformational substates) (Fig. 1.1) is not difficult. In the conditions where protein is native, these conversions are reversible and less favoured (*equilibrium intermediates*) because the protein tends to stay in the free

energy minimum. The energy landscape is modified by changing physical or chemical parameters and it is possible that the state with a minimum in the free energy, is different. In this case the protein is in a stable intermediate configuration. Lysozyme is known to stay in stabilized intermediates in some conditions [155, 215, 101].

Mean residual ellipticity at 231 *nm* shows a quite cooperative behavior in function of temperature at the lowest TFE concentration, as elicited also from correlation between ellipticity at 231 and 196 *nm* (Fig. 3.4). But at 20 % TFE the transition is not so much cooperative respect to fluorescence measurements.

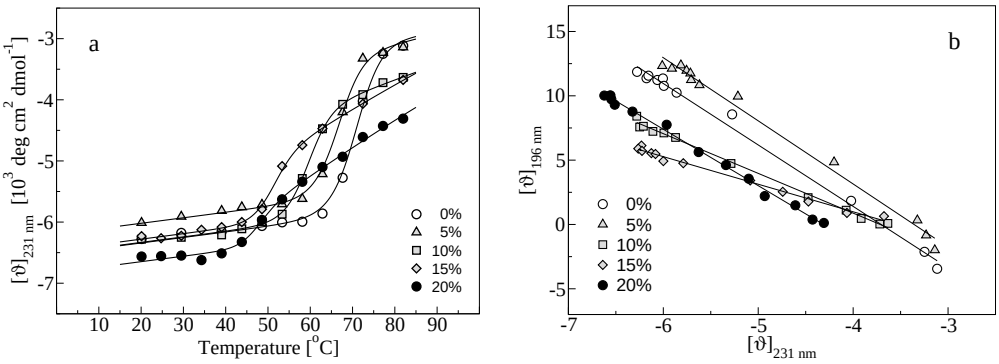

Fig. 3.4: Panel *a*: Temperature dependence of the mean residual ellipticity (θ_{231}) in lysozyme solutions at 231 *nm* for different TFE concentrations. The lines are fits to data by a three-states model (Eq. 3.1-3.2-3.3). Panel *b*: Correlation between mean residual ellipticity at 231 and 196 *nm*.

These observations suggest the existence of a third conformational state. Ellipticity at 222 *nm*, (it correlates also with signal at 209 *nm*) shows that at higher TFE concentrations, expecially at 20 %, there is a negative maximum (Fig. 3.5). Variations of CD spectra in region approximately between 209–222 *nm* is related to changes in α-helix structures. A negative increase of ellipticity at 222 *nm* was observed also at pH 2, where 15 % TFE stabilized an intermediate state of lysozyme. In order to take into account the existence of the intermediate structure, we fit the data of Fig. 3.2b, 3.4a and 3.5a with a three-states model (Eq. 3.1), describing thermal unfolding as a transition from a *native* (N) to an *unfolded* (U) state through an *intermediate* (I) state

[35]:

$$\begin{aligned} x(T) &= \left[A_N^{(x)} + S^{(x)}T\right] f_N(T) \\ &+ \left[A_I^{(x)} + S^{(x)}T\right] f_I(T) \\ &+ \left[A_U^{(x)} + S^{(x)}T\right] f_U(T) \end{aligned} \tag{3.1}$$

where $x(T)$ is one of the considered optical signals, namely the first moment of the tryptophan fluorescence emission band, the ellipticity at 231 and 222 nm, $A_N^{(x)}$, $A_I^{(x)}$ and $A_U^{(x)}$ are the intercepts related to the native, intermediate and unfolded state respectively, $S^{(x)}$ is the slope of the optical signals, which is kept the same for the three states and is zero for the PL signal and 9 $deg\ cm^2\ dmol^{-1}\ K^{-1}$ for the CD signals. Eq. 3.1 rigorously holds for the intensity of an optical signal, while in the case of the first moment of the emission spectrum it would be exact only if the quantum yields of the two states were equal. Nevertheless, it can be shown that the expression is still valid if the transition enthalpy is higher than a few $kJ\ mol^{-1}$, and the quantum yields of the two states do not differ by several orders of magnitude (D'Amico et al., Submitted to J. Phys. Chem. B). Therefore, the expression is correct in all the typical cases related to unfolding or other conformational transitions.

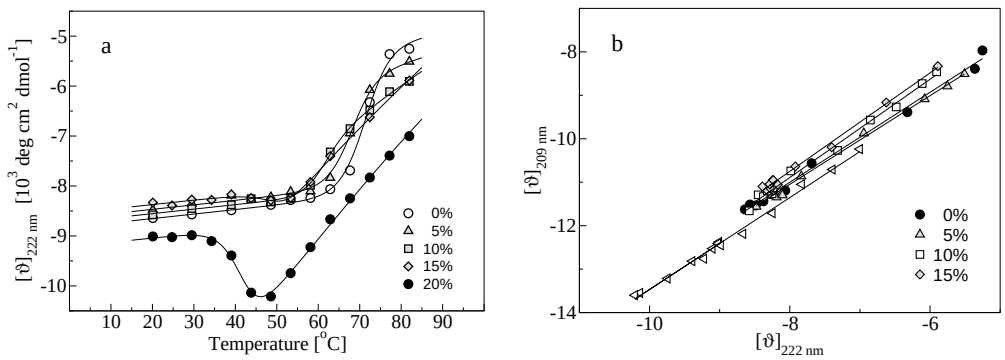

Fig. 3.5: Panel *a*: Temperature dependence of the mean residual ellipticity (θ_{222})in lysozyme solutions at 222 nm for different TFE concentrations. The lines are fits to data by a three-states model (Eq. 3.1-3.2-3.3). Panel *b*: Correlation between mean residual ellipticity at 222 and 209 nm.

The fraction of the populations of the native, intermediate and denatured

states are:

$$\begin{aligned} f_N &= Z^{-1} \\ f_I &= Z^{-1}\exp\left(-\beta\Delta G_{NI}\right) \\ f_U &= Z^{-1}\exp\left(-\beta\Delta G_{NI} - \beta\Delta G_{IU}\right) \end{aligned} \tag{3.2}$$

where $\beta = (k_B T)^{-1}$, with k_B being the Boltzmann constant, and Z is the partition function, defined so that $f_N + f_I + f_U = 1$. The free energy change between two states i and j is given by the *modified Gibbs-Helmholtz equation*, (Eq. 3.3) [23].

$$\begin{aligned} \Delta G_{ij} &= \Delta H_{ij} + \Delta c_{p,ij}(T - T_{ij}) \\ &- T\left[\Delta H_{ij}/T_{ij} + \Delta c_{p,ij} ln\left(T/T_{ij}\right)\right] \end{aligned} \tag{3.3}$$

where T_{ij} is the midpoint temperature, ΔH_{ij} is the van't Hoff enthalpy at T_{ij}, and Δc_{ij} is the difference in the heat capacity between the two states. Expressions 3.1, 3.2 and 3.3 were used to interpolate the data of Fig. 3.2b, 3.4a and 3.5a. An iterative fitting procedure was used by setting specific constraints on the parameters obtained at each iteration. The differences in the heat capacity between the native and the intermediate states and between the intermediate and the unfolded state were set respectively to $\Delta c_{p,NI} = 6.5\ kJ\ mol^{-1}\ K^{-1}$ and $\Delta c_{p,IU} = 0$ [257].

We obtained $\Delta H_{NI} = 330\, kJ\, mol^{-1}$, $\Delta H_{IU} = 24.5\, kJ\, mol^{-1}$, $T_{IU} = 74\,^{\circ}C$ for all the TFE concentrations, and T_{NI} values of 73, 68, 61, 52, 43 $^{\circ}C$ for TFE concentrations 0, 5, 10, 15, 20 % v/v respectively. In brief, the shift of the transition temperature at lower values is ascribed to the transition from the native to the intermediate state. Consistently, the cooperative transition highlighted by the PL experiments is related to the formation of the intermediate state.

By analyzing CD and fluorescence measurements by a three-states model, we concluded that the shift of the transition temperature (observed in Fig. 3.2b) at lower values is due to the transition from the native to the intermediate state. This intermediate seems the same of I_α (only α-domain is structured), already observed [155, 215]. We used the parameters obtained by three-states model to plot the stability curves of our physical system at different TFE content (Fig. 3.6).

The maximum of the stability curves moves towards higher values of ΔG and this confirms that TFE has a stabilizing effect on lysozyme conformation. On the other hand, the shift of stability curves towards lower temperatures, explain the decrease of the unfolding temperature, which can be defined as the temperature where $\Delta G = 0$.

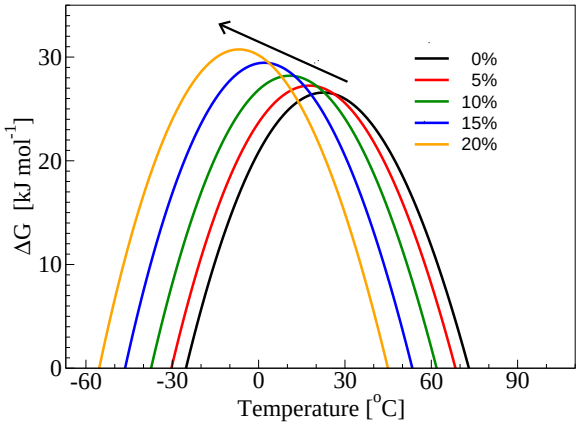

Fig. 3.6: Stability curves of lysozyme solutions at different TFE content (0, 5, 10, 15, 20 %) drawn by using the parameters obtained by fits of all experimental results (Fig. 3.2b, 3.4a and 3.5a) with a three-states model (Eq. 3.1-3.2-3.3). ΔG is the free energy between the native and the unfolded state, in presence of an intermediate.

These stability curves suggest another interesting aspect of this system, but it is out from the focus of our work. In particular stability curves across the line $\Delta G = 0$ at two temperatures: the highest characterizes the typical unfolding (an entropy-driven process), the lowest characterizes the cold denaturation (an enthalpy-driven process).

3.1.5 Enhanced aggregation of the intermediate conformer

We tested the propensity to aggregation of lysozyme samples at 20 mg/ml at different concentration of TFE upon incubation at 65 $°C$. They become turbid very fast, so a lot of aggregates are formed. However we tested the possibility that these aggregates can form supramolecular structures. Visual observations of samples (Fig. 3.7) highlight that at low (<10%) and high (>30%) TFE concentration precipitates are visible in the bottom of the cuvettes and solutions are turbid. On the contrary, in the intermediate range of TFE (between 10% and 20%) the sample is gel-like because a strong network is formed. Also, we performed Thioflavin T assays on these samples (see § A.2). Because of macroscopic network of aggregates, the sample was not homogeneous, but variable size bundles were in the cuvette. Thus we filtered the sample by sterile gauze and mixed with ThT solution at final concentration of 25 μM.

Fig. 3.7: Lysozyme solutions (20 mg/ml) in PBS buffer with increasing 15 % TFE content (0, 5, 10, 15, 20, 30, 50 % from left) after incubation at 65 $°C$ for 20 days.

Fig. 3.8 displays the PL intensity of incubated samples upon addition of Thioflavin T (25 μM). The ThT luminescence peak values were divided for the residual lysozyme and ThT concentrations measured by absorption spectroscopy. The results indicate an increasing of ThT luminescence, for increasing TFE content up to 15–20 %, whereas the ThT signal decreases at greater TFE concentration, consistently with the visual inspection of Fig. 3.7. It is worth to note here that the lowest PL value is almost three orders of magnitude greater than the PL intensity of thioflavin just added in a not incubated lysozyme sample, indicating a consistent aggregation also for the sample without TFE.

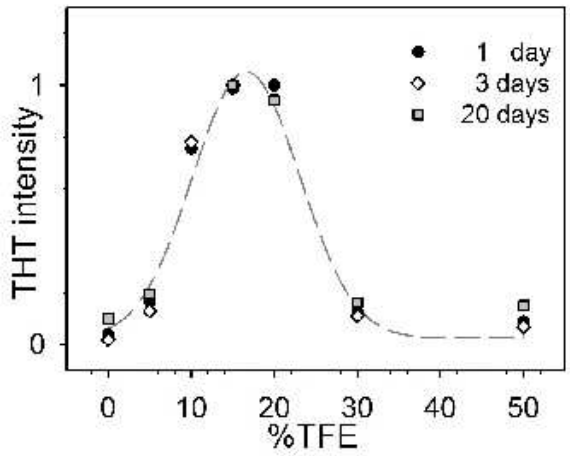

Fig. 3.8: ThT assays (25 μM) in lysozyme solutions (20 mg/ml) filtered by gauze after incubation at 65 $°C$ for 1-3-20 days. Normalized emitted light at 482 nm.

Lysozyme has been found to form amyloid fibrils in the presence of TFE upon incubation at high temperatures [130]. We performed AFM experiments in order to check if the enhanced aggregation propensity at 10–15% TFE is consistent with the formation of amyloid fibrils. We observed these samples by AFM imaging in *tapping mode*. We prepared samples on mica as described at the end of § 2.3.2, but we preliminarly filtered the samples by gauze, because of too big aggregates, and diluted one hundred times. Fig. 3.9 shows an image of a fibril on an amorphous background.

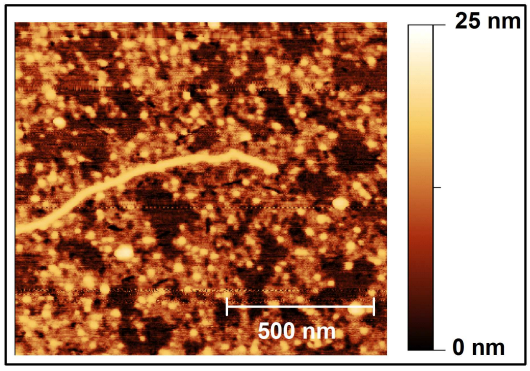

Fig. 3.9: AFM image of a lysozyme filtered sample (20 mg/ml) in PBS buffer solution with 15 % TFE and incubated at 65 $°C$ for 20 days.

This is in accord with known results from other authors [130], but the picture is poor of fibrils because of experimental difficulties. In fact it is worth to highlight that AFM imaging needs that observed sample is low concentration and we had an high concentrated one with interconnections extended on all the sample. So we had also break mechanically and gently the network before perform the measurement.

3.2 Conclusive remarks: thermodynamic stability determined by molecular conformation.

The relation between thermodynamic stability and molecular conformation is well accepted and it is a basic concept in folding process. In fact the folding pathway is driven by a progressive decrease of free energy, that is the physical parameter which characterizes the stability. Each local minimum in the energy landscape (Fig. 1.1) corresponds to one protein conformation.

When physical conditions of solution are changed the native protein moves towards another local minimum by changing its own conformation. This process is controlled by balance of intra- and intermolecular interactions and kinetics of these reactions are ruled by the height of energy barriers. In this chapter we observed the preferential hydration of Trp 62 and its environment (Fig. 3.1b). Secondary (Fig. 3.1a) and tertiary structure of lysozyme is not affected by increasing amount of TFE at room temperature, but change at high temperature (Fig. 3.2-3.3). Maybe the greater surface area exposed, due to partial unfolding, favours the contact among TFE molecules and the protein. This would enhance the effect of TFE on protein conformation. It would be interesting to study the behavior of this sample by adding a competitor of TFE and which does not affect the dielectric constant of the solvent. We also highlighted the stabilization of an unfolding intermediate around 20% TFE (Fig. 3.3-3.5). Incubation at high temperature is compatible with the presence of fibrils in solution with 15–20 % TFE (Fig. 3.8-3.9). This not monotonous behaviour with TFE concentration can be ascribed to the degree of preferential solvation of lysozyme. In fact, at low concentration, TFE molecules are too few to guarantee a relevant effect on protein conformation. On the other hand, at high concentration, TFE molecules manifest strange effects like self-clustering [207], so they prefer to interact with each others rather than with proteins. Therefore fibril formation seems strongly related to protein conformation and solvent properties. In order to investigate better this important aggregation pathway, in the next chapter we will address our study on lysozyme solutions in fibrillogenesis prone conditions.

Chapter 4

Lysozyme stability and fibrillation

4.1 Fibrillogenesis of lysozyme

4.1.1 Introduction

In the previous chapters we discussed the stability of lysozyme at neutral pH and we addressed the aggregation processes at high temperature, with or without TFE. In particular we started to study lysozyme solutions at conditions which are prone towards fibrillogenesis (Chapter 3).

In this chapter we want to examine more deeply this process that is very important in pathology and in biotechnology. Indeed fibril formation is strongly involved in several amyloidogenic deseases and neurodegenerative pathologies. The role of fibrils and protofibrils on the cell damage is not clear yet and the specific literature is abruptly increased in the last few years. Initially the observations on damaged tissues suggested that fibrils were the real cause of the deseases (*Amyloid hypothesis*) [105]. In the last decade this hypothesis was strongly criticized and now the most accepted idea is that oligomers and small precursors of fibrils are the most toxic species for the cells [185]. Different mechanisms were proposed for the citotoxicity of oligomers: they could affect biochemical processes within the cell or could permeabilize the cell membrane and lead to an uncontrolled flux of ions between the inner and the outer part of the cell. It is not clear yet if all pathological proteins form oligomers according to the same mechanism or if different proteins use different mechanisms. The effect of both fibrils and oligomers of lysozyme on cells was recently studied and the results indicate that both oligomers and fibrils are dangerous and work through different mechanisms [185]. In particular for lysozyme, fibrils damage cell membrane and lead the cell to a *necrosis*-like death. On the other hand, the oligomers activate a cascade of intracellular signals and lead to an *apoptosis*-like death in a longer time [88].

These issues stress the importance of a better knowledge of the initial step of the fibrillogenesis process, when the oligomers form, and the factors that affect the kinetics, the morphology and the stability of oligomers. We tried to reach this goal by studying lysozyme fibrillogenesis. Hen egg-white lysozyme fibrillates in different conditions, by controlling suitably pH, temperature, ionic strength and addition of cosolvents. We collected all these conditions in a summary table:

pH	T [$°C$]	buffer+cosolvents	time	ref.
7.4	37	30% TFE	day	[130]
6.9	85 $\xrightarrow{slow}$ 25	DTT 20 mM	1–2 hours (fibril gel)	[269]
6.3	50	GuHCl	day	[260]
5.5	25	acetate buffer 50 mM + TFE>18%	day	[63]
4	25	90% ethanol + 10 mM $NaCl$	day	[91]
3	57	$NaCl < 13$ mM + HCl	11 days	[7]
2	57	HCl	2 days	[7]

Tab. 4.1: Typical fibrillation conditions for hen-egg white lysozyme.

4.1.2 Fibrillogenesis of HEWL at acidic pH: the reasons

Because we are interested to study the balance of interactions in the fibrillogenesis prone lysozyme, we had to choose the most simple conditions. We addressed our study by focusing on the conditions described by Arnaudov et al. [7]. This system is simple to think about and easy to model if one wants to perform a computer simulation because it contains the protein, the water molecules and few small ions. In the other conditions described in Tab. 4.1, alcohols or more complex molecules are present, so the competition of preferential solvation by water molecules on co-solvents and on proteins is very important to characterize the system [38, 189], as we highlighted in Chapter 3. Thus the presence of these components makes very difficult the quantitative description of the system.

We also chose to study fibrillogenesis of lysozyme at acidic pH for other reasons. The most important is to investigate the role of electrostatic repulsion on the intra- and intermolecular potential. We strongly shifted the balance of interactions towards the attractive contributions by increasing the temperature, which favours the partial unfolding of the protein and can affect the aggregation pathway or its evolution rate. This is typically affected by ionic strength through the modulation of the range of action of electrostatic interaction. The increasing of the aggregation speed could turn an ordered aggregation process into an amorphous one. Second, the reproducibility of fibrillogenesis kinetics is better than other conditions and

other proteins [161, 267, 181, 88]. Also, the formed fibrils are amyloid-like as confirmed by X-ray diffraction through the typical *cross-β* pattern [162]; in particular the distance between β-strands along elongation axis is 4.7 $\mathring{A}$ and the distance between two β-filaments is 10 $\mathring{A}$. This is in accord with the observation that lysozyme fibrils are not stabilized by covalent bonds (e.g. disulfide bridges) but only by physical bonds [7].

Just for completeness, it is worth to notice that lysozyme at pH 2 and high temperature (57 $°C$ or more) can be hydrolyzed after a long incubation time, but this process is not relevant for fibril formation [7].

4.2 Lysozyme conformation at acidic pH

As we highlighted in the Chapter 1, the stability of proteins is quite shaky and even modest changes in the environment can quickly affect the protein structure and its thermodynamic characteristics. This sensitivity to environmental factors can be used to exploit proteins as biological sensors. Studies about that are growing and recently Arosio and his co-workers have engineered GFP (Green Fluorescent Protein) to monitor pH and chloride concentration *in vivo* [8].

We studied the effect of pH on the thermodynamic stability of lysozyme by comparing the excess of heat capacity at pH 7 and pH 2, the cases which we are interested on. Fresh lysozyme solutions were prepared at 37.3 mg/ml in PBS 0.1 M at pH 7 and water corrected at pH 2. Calorimetric thermograms were collected at 0.13 $°C/min$. Fig. 4.1a shows that the profile at pH 2 is peaked at very lower temperature and is more spread. This suggests that the protein is destabilized because the unfolding starts at lower temperatures (beyond 45 $°C$) than neutral pH (beyond 60 $°C$). Further, the spreadness of the profile at pH 2 suggests a greater flexibility of the protein and the presence of more conformational substates. As we will see better in the next paragraphs, lysozyme is very charged at pH 2. In particular, the protonable residues are neutral or positively charged, so the protein destabilization is due to internal electrostatical repulsions and to cleavage of surface salt bridges, which exist at pH 7 [243]. We have already stressed the importance of these bridges in the Chapter 2, about the effect of ionic strength. In this regard, we recall that at pH 7, in a salt free solution, the peak temperature and the change of enthalpy between native and unfolded state are bigger than in phosphate buffer. These observations suggest that also the change of enthalpy at pH 2 is lower than pH 7 in a salt free solution. However, the profiles in Fig. 4.1a are characterized by similar ΔH because of the difference of ionic strength in the used samples.

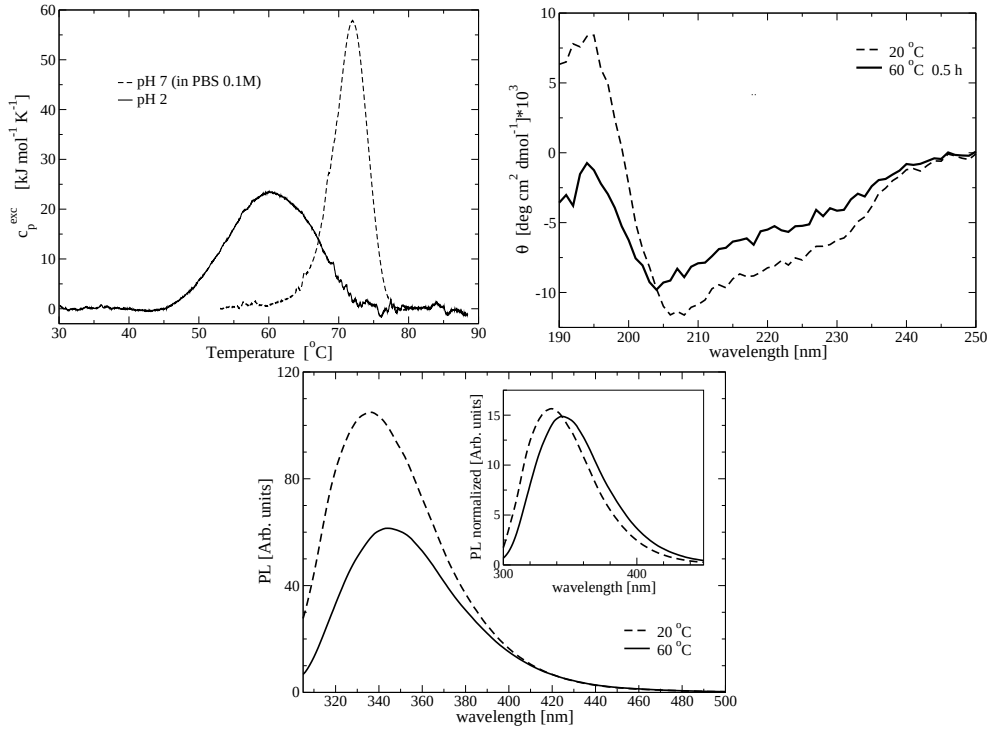

Fig. 4.1: a) Excess of heat capacity of lysozyme solutions (37.3 mg/ml) in 0.1 M phosphate buffer at pH 7 (dashed line) and in water corrected at pH 2 (solid line). Scan rate: 0.13 $°C/min$. b) Far UV CD spectra (OLIS spectrometer) of a lysozyme solution (0.21 mg/ml) at 25 (dashed line) and 60 $°C$ (solid line). c) Photoluminescence by tryptophans at room temperature (dashed line) and 60 $°C$ after thermalization (solid line) (λ_{exc}= 300 nm); inset: spectra normalized to total emitted intensity.

Afterwards we investigated the effect of the temperature rise on the secondary structure of lysozyme. We prepared lysozyme solutions at 0.21 mg/ml at pH 2 by the same procedure. We collected the far UV CD spectra at 25 and 60 $°C$, just after the thermalization of the sample. Fig. 4.1b indicates a relevant conformational change of the secondary structure, during a few tens of minutes. This is in accordance with Arnaudov' results [7]. He observed these changes in the pH range 2–4 at different temperatures between 25 and 85 $°C$. In particular, CD spectra are unchanged up to 45 $°C$ at pH 2; the spectra change between 50 and 60 $°C$ consistently with Fig. 4.1a, and coincide at higher temperatures. Ellipticity between 208 and 230 nm is less negative by increasing the temperature and the minimum at 208 nm shifts towards lower wavelengths, up to 202 nm. At pH 3 and 4 these observations are the same, except the highest variability range of CD spectra: it is

respectively 65–80 $°C$ and 70–85 $°C$. For all pH values Arnaudov observes an isodichroic point at 204 nm: this suggests a two-states cooperative process [67]. The deconvolution of spectra by *CONTINLL* (see Appendix A.8) shows a decrease of content of α-structure and the increase of β-structure and *random coil*.

We also studied the conformational change, thermally induced, on the tertiary structure through tryptophans emission (Fig. 4.1c). We used an aliquot of the same sample used for CD measurements. After thermalization of the sample, at 60 $°C$, the spectrum is strongly red-shifted and the intensity is lower than the spectrum collected at 20 $°C$. The inset shows the spectra normalized to total emitted light. Tryptophans are sensitive to polarity of the local environment and a *red-shift* of their emission band is correlated with an exposure to a more polar environment [133] (as seen in Chapter 3) . That means the protein is changing tertiary structure and is more exposed to the solvent. A similar plot is also obtained for the spectra collected at 65 and 70 $°C$.

4.3 Lysozyme stability in acidic solution

4.3.1 Isothermal compressibility and collective diffusion

We prepared a fresh stock solution of lysozyme in *Super MilliQ* water adjusted at pH 2 at 25 $°C$ by adding small amounts of hydrochloride acid. Then we filtered the sample by *Millex LG* filters. In order to be absolutely sure that only monomers were present in solution we filtered the sample again by 30 kDa filters. Afterwards we suitably diluted the stock solution by acidic water and prepared eight samples at different concentration in the range 0.8–12 mg/ml. Finally we filtered the samples directly in cuvette. We measured the scattered intensity (by SLS) and the correlation function (by DLS) at 90 ° (λ= 532 nm) and as a function of the temperature in the range 20–70 $°C$. We waited just 5 minutes to thermalize the sample at the given temperature and we hold it in the thermostated block just enough time to collect the measurements. Static light scattering experiments allowed us to build the compressibility curves (Fig. 4.2).

Osmotic isothermal compressibility χ is a thermodynamic property which can be measured by light scattering experiments (Eq. 4.1):

$$\frac{1}{nRT\chi} = \frac{M_0}{RT}\frac{\partial \Pi}{\partial c} = \frac{KcM_0}{R_{90}} \tag{4.1}$$

where Π is the osmotic pressure and R_{90} is the Rayleigh ratio at 90°

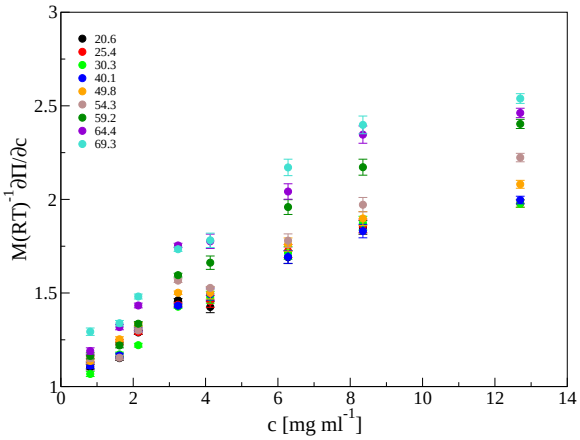

Fig. 4.2: Compressibility curves at different temperatures: pH has been adjusted for each protein concentration.

($R_{90} = \frac{I}{I_0}\frac{r^2}{V_s}$). I, I_0, r, V_s are respectively the scattered and the incoming intensity, the sample-detector distance and the *scattering volume*. In an homogeneous and isotropic system of identical particles which interact by a potential of mean force (PMF) $v(r)$, Π depends upon $v(r)$ and pair correlation function $g(r)$ (Eq. 4.2) [104]:

$$\frac{\Pi}{k_B T} = \rho + \frac{2}{3}\pi\rho^2 \int_0^\infty \frac{\partial e^{-v(r)/k_B T}}{\partial r} g(r) e^{v(r)/k_B T} r^3 \, dr \tag{4.2}$$

Correlation functions collected by dynamic light scattering experiments at 90 ° allowed us to measure the collective diffusion coefficient D_c of molecules. D_c is an useful parameter to describe brownian dynamic of interacting proteins in solution. D_c depends upon intermolecular interactions and the hydrodynamic effect [226]. Fig. 4.3 shows D_c normalized by the diffusion coefficient at infinite dilution D_0.

D_0 contains the direct effect of temperature (thermal energy increases the mean velocity of particles) and its indirect effect upon the change of bulk viscosity. The inflection of the curves at the highest concentrations is a footprint of second order concentration effects. Besides the increase of the diffusion coefficient by the concentration is an inusual effect due to long range interactions. Indeed the typical behavior is that diffusion is reduced by increasing concentration because of *crowding* and *short range* attractions. However some works show that in some systems the behavior is the reverse [230]. We can imagine, in a very simple way, that if an highly charged particle runs against to another one, they "feel" each other very long before to collide

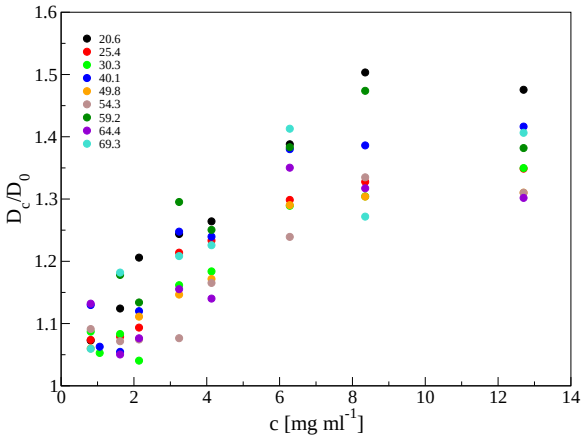

Fig. 4.3: Collective diffusion coefficient as a function of protein concentration and temperature. It has been evaluated by DLS correlation functions. It is normalized respect to infinite diluition sample.

and change their state of motion: they can deflect, reverse direction, repulse. Anyway it is difficult that they can enter into contact and suffer energy losses due to protein-protein collisions, mechanical deformations and so on.

4.3.2 Virial expansion

In dilute solutions we can expand the isothermal compressibility as a function of all powers of ρ (number concentration of protein):

$$\frac{\Pi}{k_B T} = \sum_{i=1}^{\infty} B_i \rho^i \tag{4.3}$$

where $B_1 = 1$. Second virial coefficient B_2 can be expressed by potential of mean force $v(r)$.

$$B_2 = \frac{2}{3}\pi \int_0^{\infty} \frac{\partial e^{-v(r)/k_B T}}{\partial r} r^3 \, dr \tag{4.4}$$

It is simple to proof that the following approximation is valid in diluted solutions.

$$\lim_{\rho \to 0} g(r) e^{v(r)/k_B T} = 1 \tag{4.5}$$

A negative B_2 means that intermolecular interactions are mainly attractive, a positive B_2 means that they are repulsive.

For example, in crystallization conditions, B_2 is between -1 $\cdot$ 10^{-4} and -8 $\cdot$ 10^{-4} $\frac{cm^3 mol}{g^2}$ (*crystallization slot*) [87, 214, 160, 191]. We know that it

depends upon salt content [250, 213], type of added salt [62, 73], net charge of the protein [188], polarity of the solvent [166], molecular conformation [156]. We analyzed curves in Fig. 4.2 by the following expression [24, 184]

$$\frac{KcM_0}{R_{90}} = \frac{M_0}{M_w}\left(1 + 2A_2M_wc + 3A_3M_wc^2\right) \tag{4.6}$$

where A_2 and A_3 are respectively the second and the third virial coefficient in $g^{-2}molcm^3$. B_2 is related to A_2 through the following expression:

$$B_2 = A_2\frac{M_0M_w}{N_A} \tag{4.7}$$

and has dimensions of a volume. It is strongly connected with the interparticle interaction volume. First, we computed M_w as the infinite diluition limit of $\frac{R_{90}}{Kc}$ through a linear and a parabolic fit:

$$M_w = \lim_{c\to 0}\frac{R_{90}}{Kc} \tag{4.8}$$

Due to the noise of the experimental points, we assigned the variability range of M_w (for each temperature) by varying the number of experimental points considered and using different ways to weight them in the extrapolation of M_w. Afterwards, we calculated B_2 by considering the quadratic term in c or not in the fitting procedure [66]. Fig. 4.4 shows the parabolic fit to data and the inset highlights that M_w sligthly decreases at high temperatures. Variability range for B_2 (for each temperature) was estimated likely

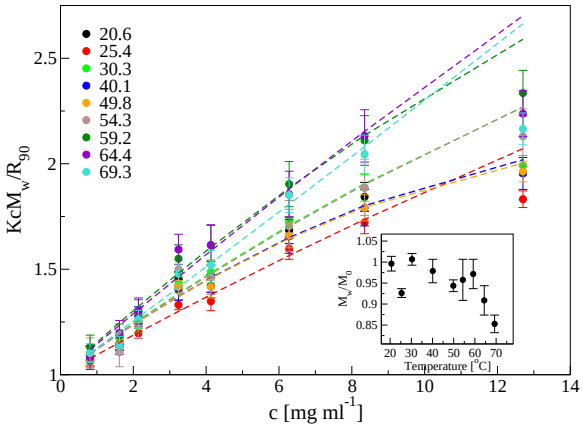

Fig. 4.4: Quadratic fit to data. The inset shows the ratio $\frac{M_w}{M_0}$ used to fit data.

M_w and using the calculated values of M_w. In particular B_2^{max} is calculated by using M_w^{max} and B_2^{min} by using M_w^{min}.

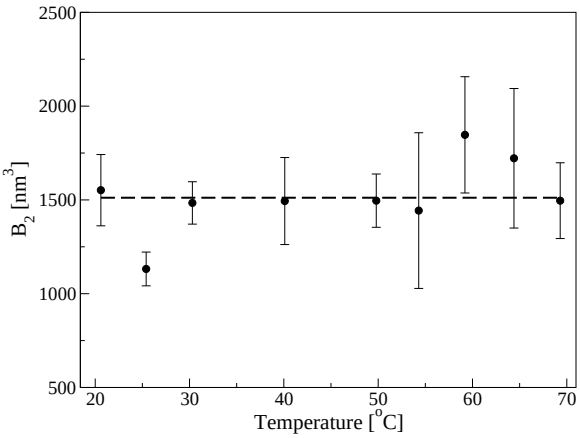

Fig. 4.5: B_2 calculated by quadratic fit in Fig. 4.4. Error bands represent the maximum variability range

Fig. 4.5 shows that changes of B_2 are within the noise level, so we can conclude that the potential of mean force is unchanged in all explored temperature range and repulsive ($B_2 \approx 1500\ nm^3$). This value is very far from *crystallization slot* (for lysozyme it would be -34 ÷ -272 nm^3) [191, 160]. A positive B_2 means that proteins are soluble and they are in a *good solvent* [13]. The *quality of solvent* is a concept introduced for polymer physics and depends upon its dielectric constant, dipole moment of its molecules and other chemical properties.

The quality of solvent, typically, decreases by diminuishing the temperature and the Θ-*temperature* is defined as the point where B_2 is zero [74]. The quality of solvent is strongly related with the propensity to form bonds protein-solvent bonds. It is based on the criterion *"like dissolves like"*. In a good solvent a protein prefers to interact with solvent molecules than another protein, so self-assembly is unfavoured. Typically a globular protein, like lysozyme, has a polar surface and an hydrophobic core. This arrangement is determined by solvent. A protein rich in hydrophobic residues or in β-sheet content prefers to stay in an apolar environment, like cellular phospholipidic double-layer or oil-based solvents.

We measured also the hydrodynamic radius of protein through dynamic light scattering measurements. The experimental physical quantity accessible is the *collective diffusion* coefficient (D_c). We have to distinguish it from the *self-diffusion* coefficient D_s. D_c is the parameter which appairs in the Fick's laws about diffusion [52]. It is a statistical process and involve collectively a large number of particles. D_c can be affected by intermolecular interactions,

hydrodynamic effects and depends by protein concentration [226].

D_s is the diffusion coefficient that characterizes the motion of a single molecule. It is concentration independent and it is related to hydrodynamic radius through the Stokes-Einstein law [24]:

$$D_s = \frac{k_B T}{6\pi\eta R_h} \tag{4.9}$$

where η is *bulk viscosity* of the solvent at temperature T and R_h is the *hydrodynamic radius* of a sphere which have the same diffusion coefficient than the protein, in laminar regime. The dependence of D_c from interactions and hydrodynamic effects is expressed by the equation (4.10) [183].

$$\frac{D_c}{D_s} = \frac{H(0)}{S(0)} \tag{4.10}$$

where $H(0) = 1 - h_2\rho$ and $S(0)^{-1} = 1 + 2B_2\rho$. $H(0)$ represents the hydrodynamic effects, $S(0)$ is the structure factor of the solution and captures the main features of interactions.

We used Eq. 4.11 to fit experimental data in Fig. 4.6, test the stability of obtained parameters and fix their maximum variability range.

$$\frac{H(0)}{S(0)} = 1 + (2B_2 - h_2)\rho + o(\rho^2) \tag{4.11}$$

where h_2 and ρ are respectively the *hydrodynamic coefficient* and protein number concentration [38].

This equation allows to find the "real" hydrodynamic radius by eliminating interactions and hydrophobic effects at the same time. In particular, by defining an apparent radius $R_a = \frac{R_c}{S(0)}$ (R_c is inversely proportional to D_c according to Eq. 4.9), Eq. 4.10–4.11 lead to the following fitting expression (Eq. 4.12):

$$\frac{1}{R_c} = \frac{1}{R_s} + \frac{2B_2 - h_2}{R_s}\rho \tag{4.12}$$

where R_s is the hydrodynamic radius of a sphere that has the same self-diffusion coefficient as protein in solution, and the values of B_2 are obtained by SLS analysis.

Fig. 4.6 shows the parabolic fit to data by using Eq. 4.12. This plot allows us to find R_s and we can also obtain h_2 (Fig. 4.7) by using the informations about B_2 found from Fig. 4.5.

The curves indicate an increase of the true hydrodynamic radius (14 %) from 1.60 to 1.82 nm in the range 20–70 $°C$. This mild swelling of lysozyme

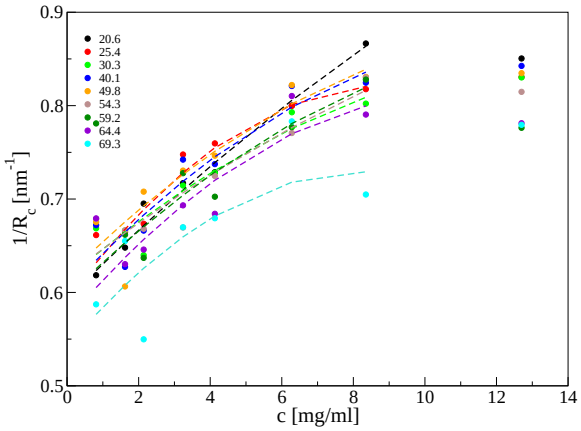

Fig. 4.6: Parabolic fit to data by using Eq. 4.12.

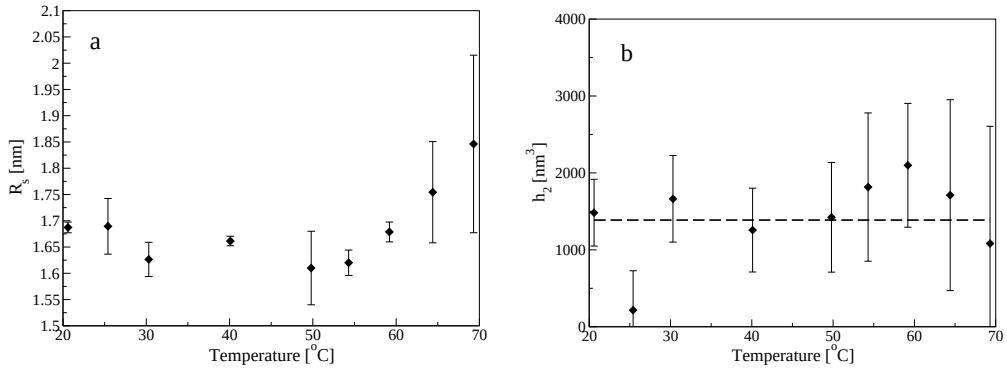

Fig. 4.7: a) R_s as obtained by quadratic fits to data by using Eq. 4.12. b) h_2 as obtained in the same way. Error bands represent the maximum variability range

starts from 60 $°C$ and R_s does not show much significative variations at lower temperatures (Fig. 4.7a). Indeed lysozyme was recently used as a diffusion tracer to measure viscosity of solutions [176]. Our result matches quite well with Arai et al. [6] who studied lysozyme solutions at pH 2.8 by SAXS. They observed no dependence upon temperature up to 65 $°C$ and an increase of R_g from 1.62 to 1.72 nm in the range 20–70 $°C$ because of changes in tertiary structure [6]. Old measurements by Nicoli et al. [171] show an increase of hydration radius of lysozyme about 18 % at pH 2 and 200 mM KCl. In particular R_s increased from 1.85 nm up to 2.2 nm by increasing temperature from 25 to 70 $°C$. By considering the change of diffusion coefficient of lysozyme solution in function of ionic strength, and the dependence of diffusion coefficient by protein charge [58], Nicoli concluded

that in our experimental conditions R_s should grow by 16 % because of thermal unfolding [171].

Anyway the values for R_s calculated by measurements in our conditions are quite far from the values that lysozyme should have if it were *random coil*. In fact, a polymer of 129 amino acids should have a gyration radius 4.4 nm [115]. This discrepance is due to the four disulfide bonds which impose some constraints to the unfolded conformations. The proof is that the reduced state of lysozyme obtained by carboxymethilation of cysteines (permanent cleavage of S-S bonds), has R_g= 3.6 nm, much closer to R_g of an ideal *random coil* polymer [115].

Fig. 4.7b shows the hydrodynamic coefficient calculated according to the same procedures used to obtain R_s. We can see that also h_2 has no dependence by temperature. Hydrodynamic interactions couple the motions of proteins from different parts of the solution through the transfer of momentum to layers of liquid between particles. Hydrodynamic interactions are relevant only at volume fractions greater than 0.01 [203] but also at low protein concentration they can be significative if long range interactions are strong. This condition can be satisfied far from isoelectric point (where protein is highly charged) and at low ionic strength (where screening effect of electrolytes is small) [230].

Competing effects keep the solution stability unchanged at different temperatures

B_2 is positive in the explored range of temperature, so repulsive interactions are prevalent. At pH 4.5, where crystallization places, B_2 is slightly negative, between zero and -8.0 $\cdot$ $10^{-4}\frac{molml}{g^2}$ [25]. The strange result is that B_2 values remain constant by the temperature, beyond 50 $°C$. Typically one expects that increasing temperature produces a larger exposure of hydrophobic residues. The hydrophobic effect pushes molecules to aggregate, so one expects that the net interaction is more and more attractive by increasing temperature. This is reasonable, but if it were true, the repulsive contribution should increase. We can immediately exclude aggregation effects (the aggregates could have a charge higher than monomers), because these DLS measurements highlight that only monomers are in solution at all the temperatures. Another reason which could affect the repulsive contribution is the protein unfolding. It has two effects: first is the increase of protein size (*swelling*); second is an eventual extra-protonation of residues which were buried. The charge of the amino acids depends upon pH of the solution and only residues which are exposed to the solvent can "feel" its electric properties. In order to evaluate how much our hypothesis about protonation is

reasonable we calculated the charge state of all residues. A protein can be represented by a sequence of amino acids, where only the aminic group of the first residue, the carboxylic group of the last one and the lateral residues of charged amino acids are protonable. The aminic and the carboxylic groups of the other amino acids cannot be included in the computation of protein charge because they are involved in the peptidic bond.

pH (Z)	POSITIVE	NEGATIVE	NEUTRAL	ENDS	Z_{max}
7 (+7.5)	11 ARG, 6 LYS	2 GLU, 7 ASP	3 TYR[†]	NH_3^+, COO^-	+9
2 (+17)	1 HIS, 8 CYS[*]	0	2 GLU, 7 ASP, 3 TYR[†]	NH_3^+, COO^-[‡]	+18[‡]

Tab. 4.2: Protonable amino acids in HEWL at room temperature. "POSITIVE", "NEGATIVE", "NEUTRAL" refer to charge of lateral residue, "ENDS" refers to charge of backbone terminals. Z_{max} is the maximum charge which lysozyme can have at a certain pH and Z is the experimental charge [212]

By comparing the true net charge (Z) of lysozyme at the two pH's with the maximum charge which can have (Z_{max}), it seems that almost all residues are exposed to the solvent at 25 $°C$. According to this calculation, only two residues could be less accessible to the solvent and the unfolding of protein could increase their exposure. Our hypothesis finds a confirm in a control

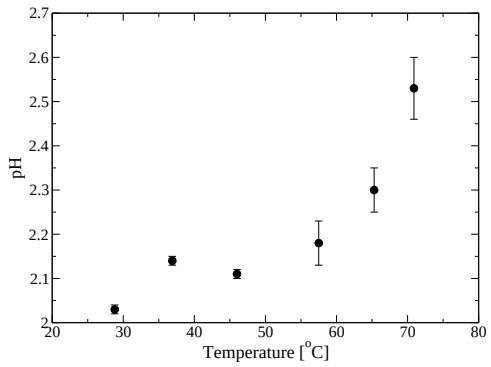

Fig. 4.8: Behavior of measured pH of a lysozyme sample at 7 mg/ml, corrected at pH 2 at 25 $°C$.

test we did. We measured the pH value of a lysozyme solution at 7 mg/ml

[*]In our case all the eight cysteines are involved in S-S bonds.

[†]They are neutral at pH lower than 10 [146].

[‡]Some old studies evidence that the unfolded state of lysozyme is fully protonated and has net charge Z_U=+19: one more charge comes from terminal carboxylic group because it changes its pK_a from $\approx$ 1.46 to 3.40 [171, 122] as a result of the unfolding.

(corrected at pH 2 at 25 $°C$) and of the solvent in function of the temperature. pH of the solvent did not increase significatively, because hydrocloride acid is a strong acid and it is disassociated at all temperatures. On the contrary, the pH of the protein solution increased almost 0.5 pH units just around 60 $°C$, the denaturation temperature at this pH (Fig. 4.8).

In § 4.3.3 we estimate the repulsive electrostatical contribution in order to make more clear the role, potentially involved, of each contribution.

4.3.3 Modelling interactions

DLVO: high electrostatic repulsion

The most commonly used potential, $DLVO$ is the sum of hard sphere, electrostatic repulsive and dispersion attractive potentials. H-S potential (Eq. 4.13) considers exclusively the impenetrability of particles:

$$W(r) = \begin{cases} \infty & r < \sigma \\ 0 & r > \sigma \end{cases} \tag{4.13}$$

where σ is the diameter of the protein. The electrostatic term takes into account the screening of elettrolytes in solution and is described by Debye-Hückel model (in M.K.S. Eq. 4.14), which is the extension of Yukawa potential for a finite size object.

$$\beta W_{el}(r) = Z^2 \frac{l_B}{r[1+\frac{a}{\lambda_D}]^2} exp\left\{-\frac{r-2a}{\lambda_D}\right\} \tag{4.14}$$

where $\beta = (k_B T)^{-1}$ and $W_{el}(r)$, a, Z, λ_D, l_B are respectively the interaction electrostatic energy of two proteins at a given distance r, the radius of the protein, its net charge, the *Debye* length (Eq. 4.15) and the *Bjerrum* length (Eq. 4.16).

$$\lambda_D = \frac{1}{\kappa_D} = \frac{1}{\sqrt{8\pi l_B N_A I}} = \frac{1}{\sqrt{8\pi l_B n_0}} \tag{4.15}$$

where κ_D, N_A, I and n_0 are respectively the *Debye* screening constant, the Avogadro's number, the ionic strength and the number concentration of ions.

$$l_B = \frac{e^2}{4\pi\epsilon k_B T} \tag{4.16}$$

where ϵ, k_B, T and e are respectively the dielectric constant of the solvent, the Boltzmann constant, the absolute temperature of the system and the electronic charge. Dispersion potentials can be described in different ways, but typically van der Waals potential is used. In order to symplify the numerical

calculations, the H-S is replaced by a rational expression (*soft-sphere* potential) and added to van der Waals term: Lennard-Jones potential (Eq. 4.17),

$$W(r) = 4u_0 \left(\frac{d^{12}}{r^{12}} - \frac{d^6}{r^6} \right) \tag{4.17}$$

where d is close to σ and represents the distance to which $W(r) = 0$. u_0 is the depth of the potential. $DLVO$ model works well for small compact proteins and at low ionic strength. It cannot describe *salting out* effects and fails at high ionic strength because it neglects the finite size of ions. We focus on the electrostatic term. Debye-Hückel model is valid only when $W_{el} << k_B T$. In our case $W_{el} \simeq 7.8\ k_B T$. So the free ions and counterions in solution have not enough thermal energy to form the electrical double layer with the typical characteristics around the colloidal charged particle (Fig. 4.9). When the protein is less charged, the ions and counterions of the double layer are in a dynamical equilibrium and are organized through the balance between the electrostatic interactions (ion-protein and ion-ion) and the entropic thermal disorder.

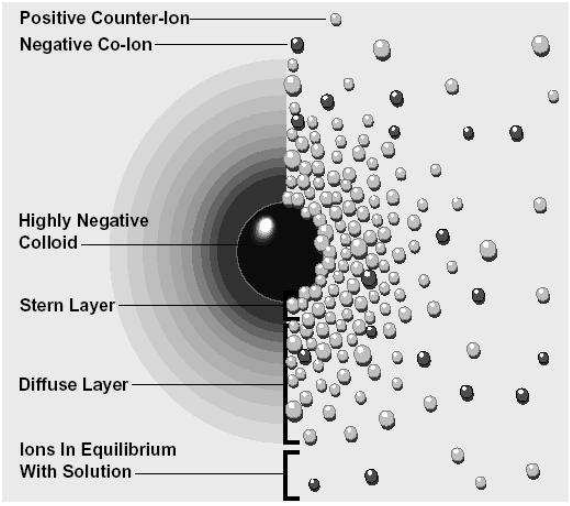

Fig. 4.9: Schematization of ionic environment around a charged colloidal particle.

We assume, as boundary conditions, that the potential is zero at infinite distance and it is finite near the particle. We describe the protein like a dielectric sphere with a surface density charge. It is surrounded by the *Stern layer*, a shell with a width about 0.2 *nm* (van der Waals radius of collapsed counterions). Beyond the *Stern layer* there is the diffuse double layer, a thicker shell. Afterwards there is the bulk solution [99]. In our

case, the electrostatic energy is so large that the counterions can collapse on the protein (*Manning condensation*). Because of the high net charge on the lysozyme surface at pH 2, we cannot use Debye-Hückel model without producing large errors.

Manning condensation

Manning condensation (or *poly-electrolite theory*) was introduced to study DNA interactions. Indeed the original theory worked only for cylindrical objects [149]. Our system can be described by a generalization of this model into spherical geometry and the calculations were explicitly reported starting since the expression for the free energy of one charged sphere in a diluted electrolytic solution [150]. The criterium for counterions condensation around colloids is that electrostatic energy between the colloid particle (e.g. a protein) and a test charge near its surface is much larger than thermal energy (Eq. 4.18) [116]:

$$W_{el} = \frac{Ze^2}{4\pi\epsilon a} >> k_B T \tag{4.18}$$

The critical surface charge density to trigger the counterion condensation, in case of large spheres, is [150]:

$$\sigma_{crit} = -\frac{e(1+\kappa_D a)ln(\kappa_D l_B)}{2\pi z l_B a} \tag{4.19}$$

where z is the valence of counterions. As a result of condensation, a certain amount of counterions collapses on the protein, so its net charge is reduced to Z_{eff} [150]:

$$Z_{eff} = \frac{\sigma_{crit}}{\sigma} Z \tag{4.20}$$

where Z_{eff}, σ and σ_{crit} are respectively net charge on the molecule after condensation, the surface charge density before condensation and the minimum surface charge density to have condensation. It is worth to note that Z_{eff} does not depend on actual Z, but only on electrolytic properties of the solvent and size of the protein.

Lysozyme can be modelled by a prolate ellipsoid (Fig. 4.10). The size of semi-axes and axial ratios are not unique in literature and the most used models are reported in Tab. 4.3. a and b are the semi-axes of ellipsoid, $\frac{a}{b}$ is their axial ratio, R_h is computed by using the expression by Tanford [244] and R_{eq}^{sph} is the equivalent radius of a sphere with the same volume.

In order to simplify further the geometry of the system, we can consider lysozyme as a sphere with an equivalent volume; so the radius must be 1.56 nm. In solution and at very low pH the tertiary structure of protein

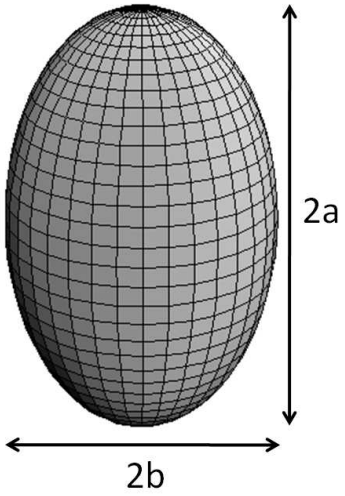

Fig. 4.10: Prolate ellipsoid.

a [nm]	$\frac{a}{b}$	R_g [nm]	R_h [nm]	R_{eq}^{sph}	ref.
2.25	1.73	1.30	1.60	1.56	[98]
2.39	1.5	1.47	1.85	1.82	[227]
2.25	1.5	1.38	1.75	1.72	[26]

Tab. 4.3: Typical parameters to model lysozyme as a prolate ellipsoid

could be a bit different than crystalline structure because of internal electrostatical repulsions. However values indicated by Guaqueta et al. [98] are in accord with our measurements.

Our experimental conditions are: $I = 10.8\ mM$, $T = 25\ ^{\circ}C$, $Z = +17$, $a = 1.56\ nm$, $S = 30.5\ nm^2$, $\sigma = 0.558\ e/nm^2$). They are useful to calculate Z_{eff} and the net charge of protein (including the counterions in the hydration shell) after *Manning* condensation.

By using the *Manning*'s formulas [150], we found $\sigma_{crit} = 0.310\ e/nm^2$ and $Z_{eff} = +9.4$. These formulas hold if $k_D a \approx 1$ and $k_D l_B << 1$. In our experimental conditions $k_D a = 0.532$ and $k_D l_B = 0.244$ (at $25\ ^{\circ}C$). Another group studied the counterion condensation in similar conditions by a different approach (not reported in a detailed way) [10]. It is valid at infinite dilution of protein and $k_D a >> 1$. However a comparison with dynamical simulations has shown a good agreement up to $k_D a \approx 1$, where Z_{eff} is slightly underestimated. By using this derivation, $Z_{eff} = +11.5e$.

We can assume that surface charge density is equal to the critical value ($\sigma_{crit} = 0.310\ e/nm^2$). That means that condensed counterions reduced net charge of the protein of $Z - Z_{eff} = 7.6$, so the number of charges collapsed on the protein must be 7.6.

Counterion condensation has two consequences. The most important is the reduction of the relative weight of electrostatic energy W_{el} with respect to thermal energy (before Manning condensation it was $W_{el} \simeq 7.8\ k_B T$; after condensation it is $W_{el} \simeq 4.3\ k_B T$). This reduction is not theoretically a

sufficient guarantee to the use of Debye-Hückel approximation. In fact, this approximation tends to overestimate the intermolecular interaction potential, especially near the protein [2]. However Alexander et al. demonstrated that even if surface potential energy W_{el} is a bit bigger than k_BT, the interaction potential retains the Yukawa form [2]. Besides, Woodward et al. [268] showed that the Debye-Hückel approximation with an effective net charge describes the experimental structure factors of very big charged objects quite well, up to 100 mM ionic strength. We can use reliably this result because they relate their results to the counterion condensation. It is worth to highlight that they did not establish quantitatively a relation between the effective charge found by condensation theory and the net charge found by their simulations. Thus, it would be interesting to know if these two quantities are the same.

The second consequence of condensation is a slight reduction of bulk ionic strength because the number of free charges in bulk solution is less. This effect, typically is negligible, but a low ionic strength, as in our case, it can be quite relevant and depends upon protein concentration. Since we are interested to explain the behavior of B_2 by temperature, we are focused on lowest concentrations samples, where the corrected ionic strength is around 10.2 mM (before Manning condensation it was $I = 10.8\ mM$). In conclusion, using Debye-Hückel approximation with the effective net charge Z_{eff}, found through *Manning* condensation theory, could be a quite good solution.

We will calculate B_2 according Eq. 4.21 [38]:

$$B_2 = 4V_I + 2\pi \int_{\sigma}^{\infty} (1 - e^{-\beta[W(r) - W_{hs}(r)]}) r^2\, dr \tag{4.21}$$

where V_I, σ, $W(r)$ and $W_{hs}(r)$ are respectively the effective interaction volume, the equivalent hard sphere diameter, the mean force potential and the hard-sphere potential. We need to estimate V_I before calculating B_2. The correct way to do that is to use the following formula [38]:

$$V_I = f_s^3(\nu_0 + \delta\nu_w)M_0N_A^{-1} \tag{4.22}$$

where f_s, ν_0, δ, ν_w, M_0 are respectively the shape correction factor, the dry protein specific volume, the weight of the hydration solvent per weight of protein, the solvent specific volume and the molecular mass. However δ is affected by a large error, so the calculated value of V_I would not be accurate. In alternative we can use $2R_h$ as a good estimate of σ and $V_I = \frac{\pi}{6}\sigma^3$ [38]. This procedure is correct for objects that have a form not so much different from spherical one. In fact Isihara [118] calculated the interaction radius R_i related to interaction volume and it is very close to function which describes the behavior of hydrodynamic radius for a rotational ellipsoid [244], provided

axial ratio not be too big. Fig. 4.11 shows a comparison between experimental and simulated values for B_2.

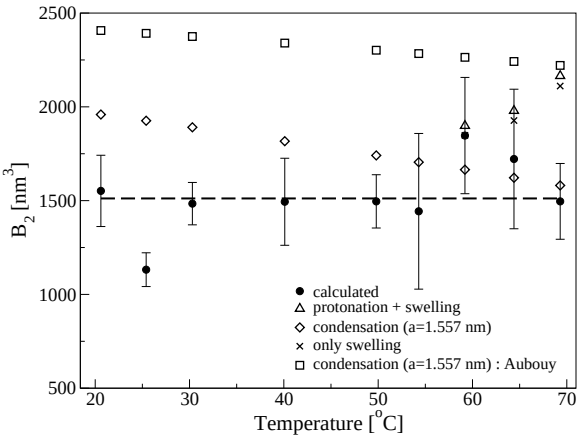

Fig. 4.11: Comparison between the second virial coefficient calculated (Fig. 4.5) from experimental data (full circles) and that simulated by taking into account four cases: no swelling of protein computed by Manning et al. [150] (diamonds), swelling (crosses), swelling and extra-protonation ($Z = +19$) (triangle up), no swelling computed by Aubouy et al. [10] (squares).

In particular we reported the second virial coefficient calculated by assuming that mean force potential is described only by a Debye-Hückel potential with the *Manning* correction on protein charge. We considered four cases. In the first case, we used 1.56 *nm* as radius of equivalent sphere [98] and computed the protein effective charge by *Manning* theory. We note a decrease of B_2 values: it is mainly due to the decrease of dielectric constant, which makes more efficient the condensation of counterions. In the second case we considered the effective charge as computed according to the formulas obtained by Aubouy et al. [10]. Because the first case fits better the B_2 values (obtained from fits in Fig. 4.4 through Eq. 4.6), we revised the first case at high temperature by supposing the *swelling* of lysozyme (third case). In particular we used the radius of the equivalent sphere [244] as computed by hydrodynamic radius obtained in our experimental measurements (1.65, 1.70, 1.82 *nm* respectively for 59.2, 64.4, 69.3 $°C$). We observe an increase of simulated B_2 as a function of the temperature. In the fourth case we further improved our modeling at high temperature by considering an *extra-protonation* (bare charge $Z = +19$ as guessed by Nicoli et al. [171]) in addition to contributions used in the third case. B_2 is not affected by extra-protonation, shows an increasing effect proportional to the protein size and it is well estimated by Manning model [150]. Of course B_2 is strongly affected by effective charge

Z_{eff} of the protein. Z_{eff} is a bit underestimated by Aubouy et al. [10] at high Z (in comparison with $NLPB$ simulations) in case of $k_D a \approx 1$; however the difference is negligible for our Z. Thus we expect that notwithstanding the underestimation of this model is a bit stronger for $k_D a = 0.532$ at high Z, it is irrilevant for $Z = +17$. So the model proposed by Aubouy et al. [10] is an upper limit for the Z_{eff} and we can consider the value calculated by Manning et al. [150] as a lower limit. Discrepancy with experimental data is easily covered if we consider an attractive contribution, modelled as a square well with $u_0 \approx 8\ k_B T$ and $\lambda = 1.2$ (Eq. 4.23)

$$W(r) = \begin{cases} \infty & r < \sigma \\ -u_0 & \sigma < r < \lambda\sigma \\ 0 & r > \lambda\sigma \end{cases} \tag{4.23}$$

with width of wall equals to $(\lambda - 1)\sigma$. Anyway we demonstrated that the electrostatic contribution is quantitatively sufficient to explain the magnitude order of experimental B_2.

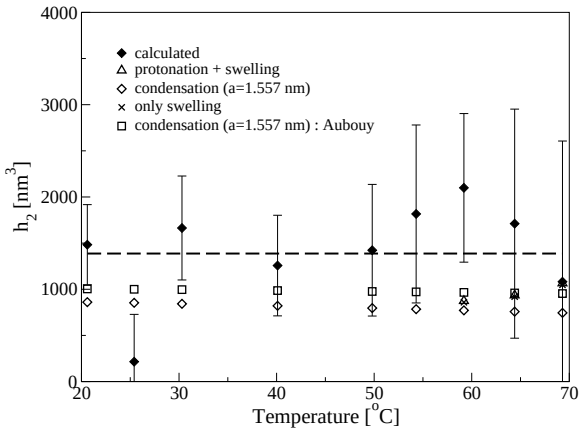

Fig. 4.12: Comparison between the hydrodynamic coefficient calculated (Fig. 4.7b) from experimental data (full diamonds) and that one simulated by assuming that the mean force potential is described only by a Debye-Hückel potential with the Manning correction on protein charge. We considered four cases: no swelling of protein (diamonds), swelling (crosses), swelling and extra-protonation ($Z = +19$) (triangle up), no swelling computed by Aubouy et al. (squares) [10].

Fig. 4.12 shows the same comparison for the hydrodynamic coefficient h_2, but the different models lead to similar results.

Simulative study through NLPB equation

The analysis performed until now assumes lysozyme as a sphere and neglects anysotropy effects or dipole moment. Besides we noted a big difference between the models for Manning condensation (diamonds and squares in Fig. 4.11). In order to study more deeply the interaction potential without particular assumptions for the effective charge, the only possible method is numeric. Several systems are studied by Monte Carlo simulations or by solving numerically the *Non Linear Poisson-Boltzmann* equation (Eq. 4.24) [94]:

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \psi(\mathbf{r})] + 4\pi \left[\rho^f(\mathbf{r}) + \sum_i e z_i c_i^b \lambda_i(\mathbf{r}) \exp[-z_i e \psi(\mathbf{r}) / K_B T] \right] = 0 \quad (4.24)$$

where $\epsilon(\mathbf{r})$, $\psi(\mathbf{r})$, $\rho^f(\mathbf{r})$, $\lambda_i(\mathbf{r})$, e, z_i, c_i^b are respectively the dielectric constant, the total electrostatic potential, the given fixed charge distribution, the accessibility parameter (equal to 0 in the region forbidden to ions and 1 elsewhere), unit electronic charge, valence of i^{th} ion and concentration of i^{th} ion in the bulk (where by definition $\psi(\mathbf{r}) = 0$). In spite of Monte Carlo methods, $NLPB$ equation does not consider the size of ions [143]. We want to know if charge distribution have some particular symmetry that could affect the intermolecular interactions. The counterions we have in solution are chloride, hydroxide and carbonate. We expect that carbonate ions have no contribution in the collapse because they are fewer than the others. Furthermore the number of collapsed ions per protein is about 8 $(Z - Z_{eff})$, as we calculated before, so the excluded volume of ions is not large and we do not need to consider the size of ions by the *Modified Poisson-Boltzmann* equation (MPB) [143]. We used 1AKI PDB file to perform the simulations. This $3D$ structure is obtained from an orthorhombic crystal of hen egg-white lysozyme, formed at pH 4.48 at 22 $^\circ C$. We used *PROPKA* software [140, 20] to calculate the correct ionization state of each residue at a fixed pH. Afterwards, we run *APBS* software [14] to solve numerically the *NLPB* equation and find the electrostatic potential of the protein. In order to do that we fixed a box large 129x129x129 $\mathring{A}^3$ and added salts to guarantee the electroneutrality of solution and the correct ionic strength. I chose to run simulations in *mdh mode* to correct some effects of the finite size of simulated system. I took care to consider the variation of the dielectric constant upon temperature. An important parameter that we had to choose was the dielectric constant of the protein. Many authors suggest to use values between 2 and 20 [86], but preliminary simulations (data not shown) highlighted a big difference between the two extreme values. A japanese group [168] studied this problem

and noted that simulations accord better with the experimental data if $\epsilon = 2$. We performed several simulations by changing pH (in the range 2–7), ionic strength I (10 mM or 100 mM) and temperature T (in the range 25–80 $°C$).

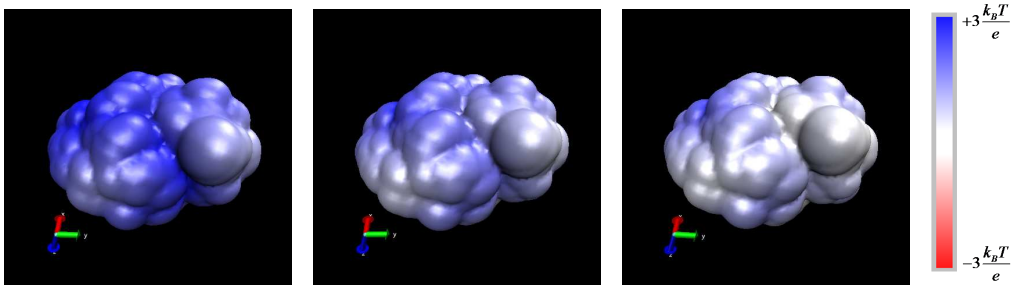

Fig. 4.13: Mapping of electrostatic potential on a surface at 5 $\AA$ from solvent accessible protein surface at $T = 60\ °C$, $I = 10\ mM$ and a) pH 2 ($Z = +17$); b) pH 4 ($Z = +8.5$); c) pH 7 ($Z = +7.5$). The false color scale runs from -3 $\frac{k_B T}{e}$ (red) to +3 $\frac{k_B T}{e}$ (blue).

We visualized the $3D$ potential generated by lysozyme in each condition and mapped it on a surface around protein at 5 $\AA$ distance, as shown in Fig. 4.13. We plotted a surface at a certain distance to understand how the proteins "see" each other when they are quite far. In fact we performed compressibility measurements at low protein concentration, so average distances are quite large.

The protein exhibits several interesting changes in potential by changing the conditions, but the potential does not have particular symmetry features. In order to represent these informations in an absolute and more simple way, we spherically averaged the potential (Fig. 4.14).

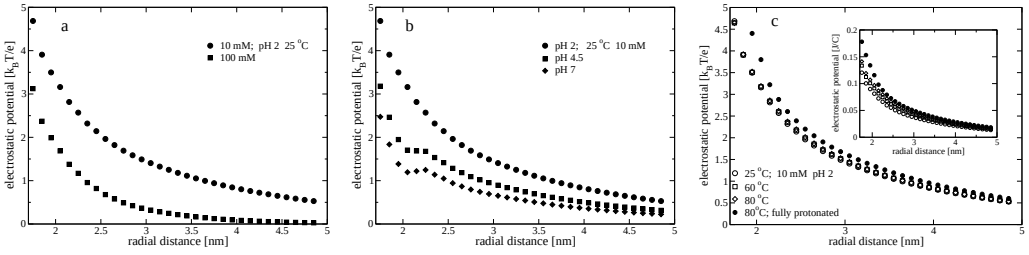

Fig. 4.14: Spherically averaged electrostatic potential of lysozyme at: *a*) different ionic strength ($T = 25\ °C$, pH 2); *b*) different pH ($T = 25\ °C$, $I = 10\ mM$); *c*) different temperatures (pH 2, $I = 10\ mM$).

As a check we fitted the potentials by a Yukawa form and they work: that is trivial since the solution was obtained by solving the P-B equation.

Fig. 4.14a shows that the potential decreases by increasing the ionic strength. This is due to the increase in the efficiency of screening. Fig. 4.14b highlights that the potential decreases by increasing the pH because isoelectric point (pI) of lysozyme is 11.2, so its net charge decreases at higher pH. Fig. 4.14c shows a negligible dependence of potential upon temperature. These curves are plot in $\frac{k_B T}{e}$ units, which is strictly related to intermolecular interactions. B_2 depends upon $\frac{\phi(r)}{k_B T}$ and even if we cannot relate $\phi(r)$ directly to B_2 ($\phi(r)$ considers the interaction protein-point charge, not protein-protein!), no temperature dependence in $\frac{\phi(r)}{k_B T}$ suggests that interactions are quite independent upon temperature. The inset in Fig. 4.14c shows the temperature dependence of electric potential in absolute units. We tried also to test the change of the net charge of lysozyme to simulate an eventual protonation of just two residues, due to unfolding (see § 4.3.2). In order to do that we repeated the simulations at pH 1, when *PROPKA* software fully protonates lysozyme. Fig. 4.14c highlights that the potential increases and this effect is much more than temperature variation. However our simulation does not consider the protein unfolding and a complete relaxation of the protein, so potentials give us only qualitative informations about the role of different parameters (temperature, ionic screening, pH) without consider geometrical variations of the protein. For this reason this software is not useful to account for eventual variations of dipole moment caused by conformational changes.

4.4 Beyond competitive interactions: the kinetics of conformational changes accompanying aggregation

We have already elicited the fast conformational changes due to the increase of temperature (§ 4.2). Arnaudov et al. observed slow conformational changes of lysozyme at level of secondary structure by incubating at high temperature [7]. Thus the following experiments point to study more deeply this process by using lower concentration samples to try to prevent the aggregation. We also tried to highlight differences due to a different temperature, around the range interesting for fibrillation.

We measured far UV CD spectra of lysozyme samples ($c = 0.21\ mg/ml$) at pH 2 at the three different temperatures as a function of the incubation time. Fig. 4.15 shows the evolution of CD spectra of the sample at 60 $°C$. The figure highlights a slow conformational change which takes a few days. An isosbestic point is evident around 205 nm and suggests a conversion from α-helix to *random coil* (compare with the cross point in Fig. A.11). A

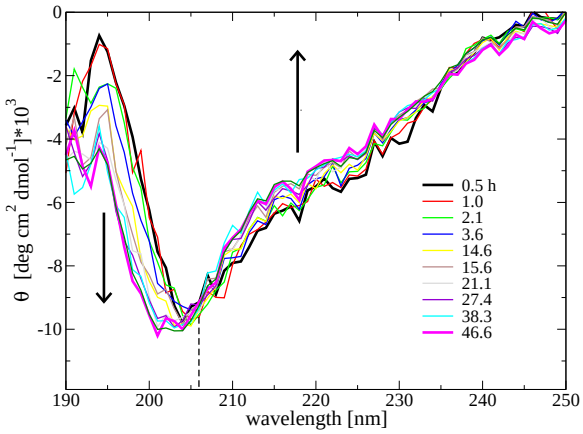

Fig. 4.15: Kinetics of CD spectra (mean residual ellipticity) at 60 $°C$. Dashed line indicates the isosbestic point.

more detailed analysis, (see §A.8) elicited that α-structure is converted both in β-structure and *random coil* (Fig. 4.16). A similar behavior is observed at 65 $°C$. At 70 $°C$ the kinetics of spectra is affected by precipitation and there are no isosbestic points, so the conformational changes are qualitatively different.

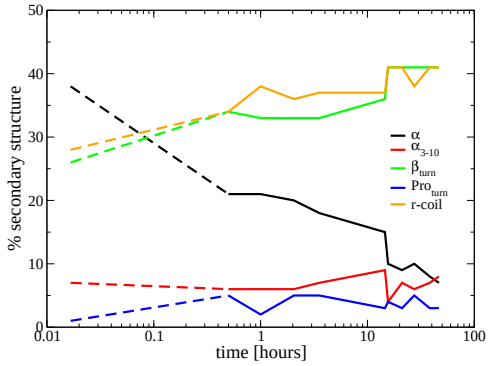

Fig. 4.16: Percentual of secondary structure as obtained from deconvolution of CD spectra in Fig. 4.15. The dashed lines indicate the change of secondary structure during the transition 20–60 $°C$.

We investigated if a slow conformational change also concerns the tertiary structure. We prepared fresh lysozyme solutions at pH 2 and 0.18 mg/ml. We used 1 cm square cell thermalized in the holder by a recirculating bath. First we measured emission spectrum of tryptophans at 20 $°C$ ($\lambda_{exc} = 300\,nm$,

$BW_{em} = BW_{exc} = 3\ nm$) and monitored the time evolution at 60-65-70 $°C$.

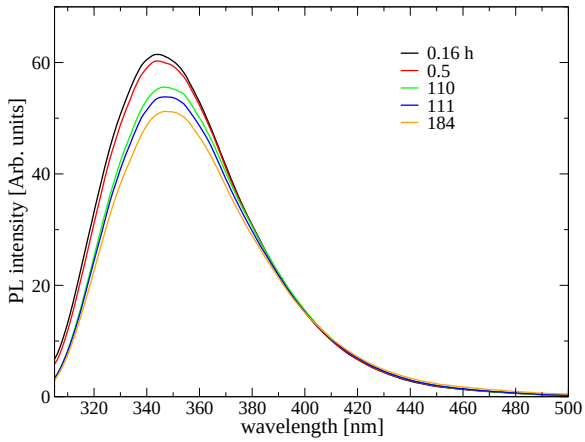

Fig. 4.17: Trp emission from lysozyme solution: kinetics of raw spectra at 60 $°C$.

As we noted in Fig. 4.1c, first spectrum collected at 60 $°C$, after thermalization of the sample, is strongly red-shifted and the intensity is lower than the spectrum collected at 20 $°C$. Fig. 4.17 highlights that this behavior goes on slightly during the incubation at high temperature and the same general trend is also observed for incubated samples at other temperatures.

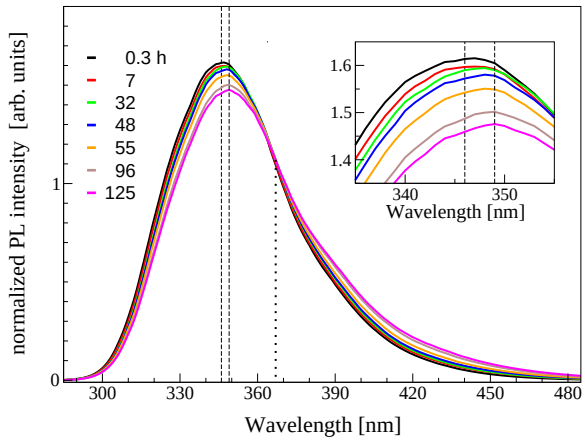

Fig. 4.18: Kinetics of normalized Trp emission from lysozyme solution at 65 $°C$. At 367 nm there is an isosbestic point. Inset: zoom on the peak red-shift.

We want to understand the origin of this red-shift and if there are some differences in the temporal behavior at these three temperatures. PL spectra

were analyzed by the following procedure. First, we eliminated the Rayleigh peak by a gaussian extrapolation. At high wavelength the spectra are less reliable because the signal is low and sometimes the buffer subtraction is not very good. Thus, we extrapolated also the signal in this range by a gaussian profile. Afterwards we normalized all the curves by their total area, that is the light emitted from the tryptophans. Fig. 4.18 shows some selected spectra at 65 $°C$ and clearly highlights the red-shift of the peak about 3 nm (see inset). In this picture we can observe also an isosbestic point at 367 nm. This point is very well defined (without considering the first spectrum). An isosbestic point typically suggests a conversion between two populations of objects.

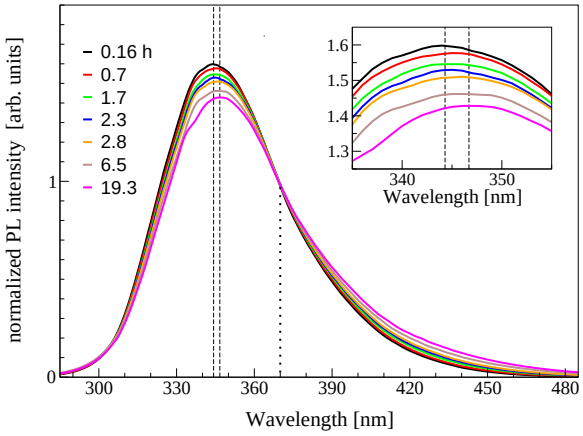

Fig. 4.19: Kinetics of normalized Trp emission from lysozyme solution at 70 $°C$. At 370 nm there is an isosbestic point. Inset: zoom on the peak red-shift.

A similar plot is obtained also for the spectra collected at 60 $°C$ and 70 $°C$. Fig. 4.19 shows, in particular, the spectra at 70 $°C$, where the peak shifts about 2.5 nm during the observation time. The isosbestic point is at 370 nm and is well defined even considering the first spectrum.

We have to choose a feature to be used as a representative parameter to plot the temporal behavior. If we used the band peak we would neglect changes in the band profile, but this seems an important feature as suggested from isosbestic point. So we chose to use the first moment of the emission spectrum as a more general parameter to describe the behavior of the emission band in function of the incubation time at 60, 65 and 70 $°C$ (Fig. 4.20).

We used *FITYK* software to fit emission spectra through three gaussians

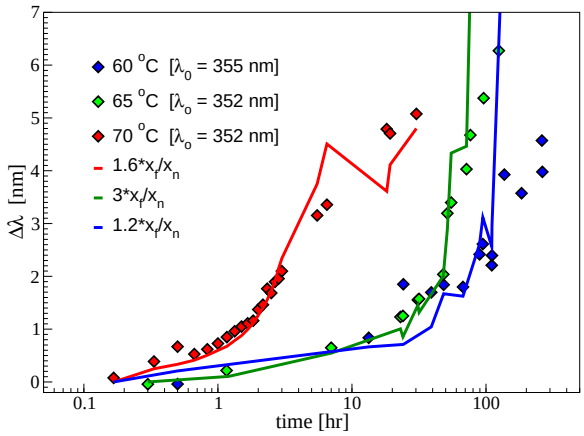

Fig. 4.20: Diamonds represent the relative shift of the first moment in function of the time at 60 (blue), 65 (green) and 70 $°C$ (red). Solid lines represent scaled ratios between final and initial fraction of molecules.

with fixed peak: 340 355 382 495 nm. We calculated the first moment as:

$$\frac{\sum(A_n X_n)}{\sum(A_n)}$$

where A_n and X_n are respectively the area and the center of each gaussian. We excluded from calculation the band at 495 nm because it is mainly due to a not reliable subtraction of buffer. We note that at the onset of kinetics at 60 $°C$, the first moment is 355 nm, that is 3 nm more than values at 65 and 70 $°C$. The reason of such a value for the first moment is that at 60 $°C$ the band is larger at high wavelengths. While the absolute value may be affected by small artifact of the spectrum, the wavelength shift is a robust parameter related to the actual conformational changes. We can note that the behavior at 60 and 65 $°C$ is initially the same until two days (the duration of the lag-phase as highlighted by Arnaudov et al. [7]). Afterwards, the rate of growth becomes higher for both kinetics. So it seems that there are two processes characterized by two different rates. We guessed that the red-shift highlighted in the Fig. 4.18 was due to an internal conformational change. But the presence of two rates in the temporal behavior of first moment suggests that there is also another origin for the red-shift and for the increase of profile width. That could be due to a conversion between a partially unfolded state (obtained during first two days) and an aggregated state. A confirm of that is the plateau at 60 and 65 $°C$, as we can see in the Fig. 4.21. Solid lines in Fig. 4.20 represent the scaled ratios between final and initial fraction of incubated lysozyme. We estimated these fractions by

fitting the emission spectra through a basis made of two spectra: the first one and the last one.

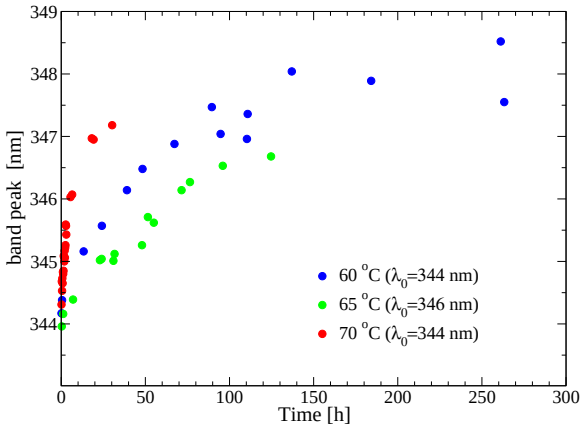

Fig. 4.21: Kinetics of band peak in function of the time at 60 (blue), 65 (green) and 70 $°C$ (red).

So the increase of the first moment, after two days, is more strongly related to the increase of profile width than red-shift of the peak. We just note that at 60 $°C$ the first moment reaches a smaller plateau, maybe because the smaller temperature changes the equilibrium between partially unfolded and aggregated states. At 70 $°C$ the first moment has an initial slope definitively higher than kinetics at lower temperatures. This observation and the well defined isosbestic point since the beginning (by using also the first spectra) suggests that at 70 $°C$ the conformational change forward to a partially unfolded state is faster and almost the total kinetics is characterized by the conversion between partially unfolded and aggregated state.

The analysis performed suggests the following explanation about the energy landscape (Fig. 4.22).

At 60 $°C$ the N^* state (protein just thermalized at high temperature; it is like an excited native state) is more stable than I (intermediate) and the energy barrier is quite high so the conformational change is slow. By increasing temperature the conformational changes are faster and at 70 $°C$, I is more stable than N^*, so the conformational equilibrium is shifted towards the intermediate state. Besides, at high concentration, protein is pushed towards the aggregated state through the funnel pathway. So we could conclude that the behavior at 60 and 65 $°C$ is initially the same, but when formation of aggregates (amorphous or fibrils) starts the sample becomes heterogeneous and the emission profile becomes more spread and Fig. 4.21 shows the difference between 60 and 65 $°C$. We guess that first moment, in the 65 $°C$

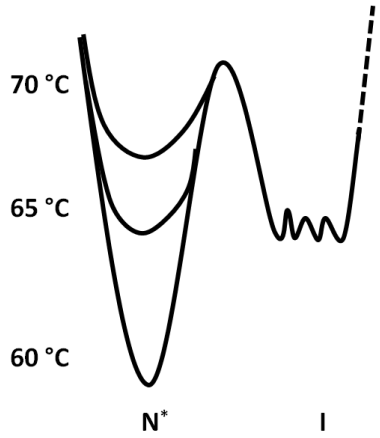

Fig. 4.22: Hypothesis of energy landscape for lysozyme sample at pH 2 with no added salt, at various temperatures.

set (Fig. 4.20), continues to increase because of additional conformational changes induced by aggregation.

By summarizing, our conclusions about fluorescence and CD data are:

- Conformational changes rapidly occur when the temperature rises from 20 $°C$ to 60-65-70 $°C$ in both secondary and tertiary structure. In particular we can observe a global exposure of residues to the solvent and, in particular, a conversion from α-helix to *random coil* and β-turn (§ 4.2);

- further conformational change slowly proceeds at high temperature as a function of incubation time. In particular fluorescence suggests a conversion from N^* to I state, CD suggests a further conversion from α-helix to β-sheet and *random coil*;

- differences in the conformational change occur as a function of incubation temperature: fluorescence measurements highlight differences among the three temperatures, so we can conclude that the effect of temperature is to accelerate tertiary conformational changes;

- conformational changes of secondary structure are less dependent upon temperature with respect to tertiary structure, but a different behavior at 70 $°C$ is also evident;

- After two days some aggregation effects are evident: Trp emission spectra show the increase of width, CD ellipticity shows the reduction of spectra because of precipitation of sample.

4.5 Kinetics of fibril formation

We prepared lysozyme samples in water at pH 2 at concentration 18.5 mg/ml. After a few minutes in the thermalized holder of light scattering equipment, we collected the correlation functions at 90° by acquiring scattered photons for five minutes. We studied the fibrillogenesis kinetics of lysozyme at three different temperatures (60-65-70 $°C$) and tried to highlight the differences among these kinetics.

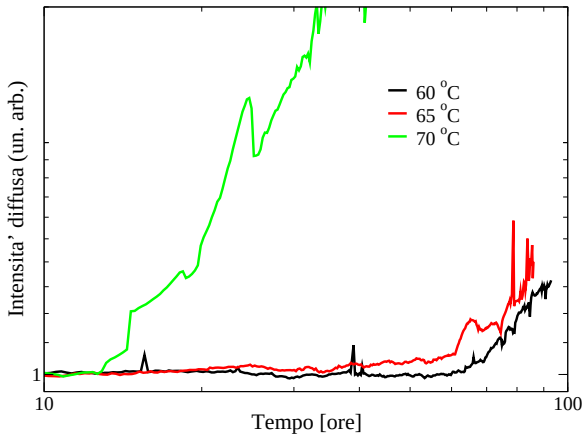

Fig. 4.23: Normalized Rayleigh ratio of a lysozyme sample (18.5 mg/ml) incubated at 60-65-70 $°C$.

Kinetics are characterized by lag-phase and an increase of scattered light. Our system is known to have some spread in the duration of lag-phase [7], so the slight difference in the duration at 60 and 65 $°C$ is not significative. On the contrary, a very shorter lag-phase at 70 $°C$ suggests that some different process is on. The growing rate seems different for kinetics at 60 and 65 $°C$, but the quite high noise level of kinetic profile makes difficult to establish a significative difference. At 70 $°C$ one clearly observes the steep increase of the intensity. A preliminary analysis of the correlation functions suggests that samples are characterized by polidispersity and some inhomogeneities. We analized the correlation functions by using *CONTIN*, an algorithm of regularization which calculates the inverse Laplace transform, that is the distribution of diffusion times [198, 199] (see §A.6). We can obtain an estimate of the size distribution of the objects in the solution by using the Stokes-Einstein law (Eq. 4.9). These distributions clearly reveal the inhomogeneities which generate the temporal sequence of the scattered intensity (Fig. 4.24).

An accurate analysis of these distributions at 60 $°C$, by focusing on the size-ranges which are always populated by objects, allowed us to highlight three main species in solution: monomers, oligomers and aggregates. We established the following ranges for the species: 1–8 nm for monomers, 8–60 nm for oligomers, 60–1000 nm for fibrils, 1–100 μm for eventual bigger objects (Fig. 4.24).

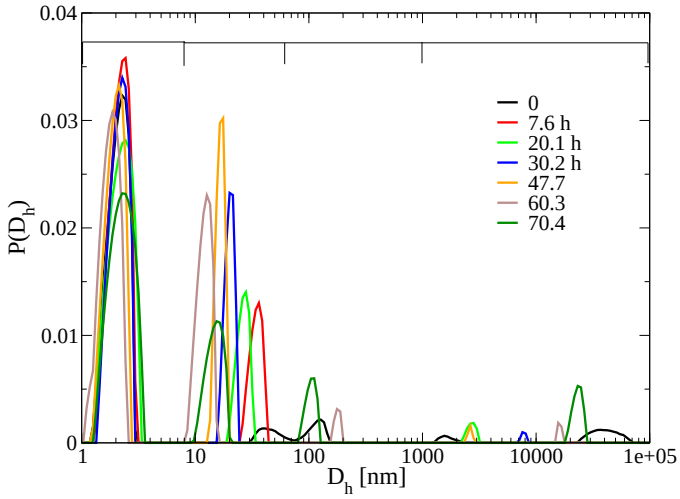

Fig. 4.24: Time evolution of size distributions in a lysozyme solution (18.5 mg/ml) incubated at 60 $°C$ at pH 2. The rule on the top indicates the selected range for $CONTIN$ analysis of size distributions.

The result of the analysis is summarized in Fig. 4.25. The size of monomers and oligomers does not change during the kinetics, so we can say that the increase of Rayleigh ratio of the oligomers is due to the increase of their number. Fig. 4.25 shows that the number of monomers slightly decreases, the number of oligomers increases and reaches a plateau during the first day of incubation. Also the aggregates appair after two days, as Arnaudov et al. already observed [7]. The species in solution have a similar kinetics at 65 $°C$. The same analysis was performed on the correlation functions collected for the sample at 70 $°C$, and it shows some interesting differences. The number of oligomers is lower, there is a large-size species that starts to grow immediately and a fourth species that seems to appair after one day of incubation.

Thus this elucidates that during the lag-phase the kinetics is characterized by the formation of large aggregates and the process involved seems quite different in times and in features than 60 and 65 $°C$.

Arnaudov et al. observed amorphous aggregates formed at the onset of

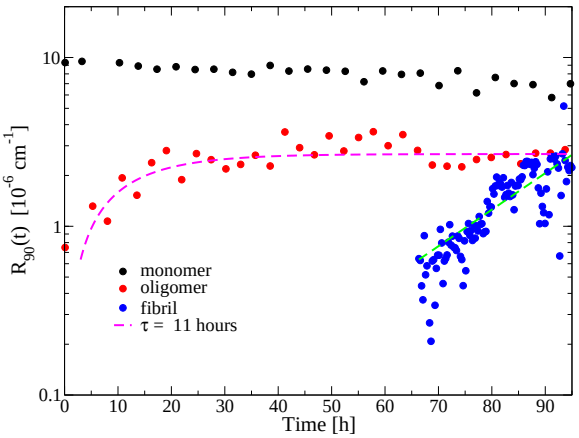

Fig. 4.25: Rayleigh ratio due to species in solution and hydrodynamic radius of the fibrils. Incubation temperature: 60 $°C$. The dashed line are guides for the eyes.

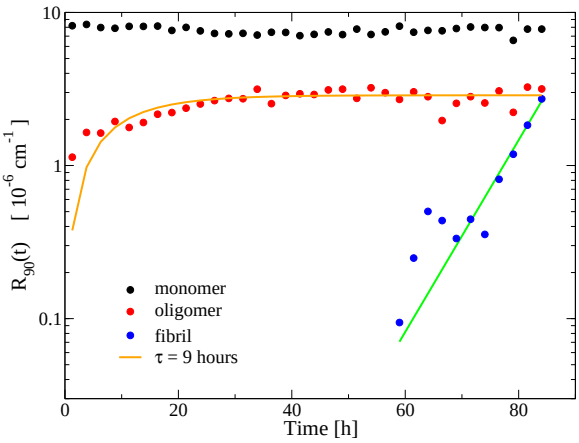

Fig. 4.26: Rayleigh ratio due to species in solution and hydrodynamic radius of the fibrils. Incubation temperature 65 $°C$. Solid lines are guides for the eyes.

incubation at 80 $°C$ by AFM; besides fibrils are formed after two days of incubation [7]. We observed this double class of objects by DLS.

On summary, CD and fluorescence kinetics reveal conformational changes at secondary and tertiary level during the incubation. At 60 and 65 $°C$ we observe the formation of oligomers which goes to completion within one day. However the biggest aggregates are visible only after two days. The most simple explanation is that large aggregates are in a small number and are not accurately revealed by DLS technique until they are more numerous. Another explanation could be that a further conformational change of the oligomer

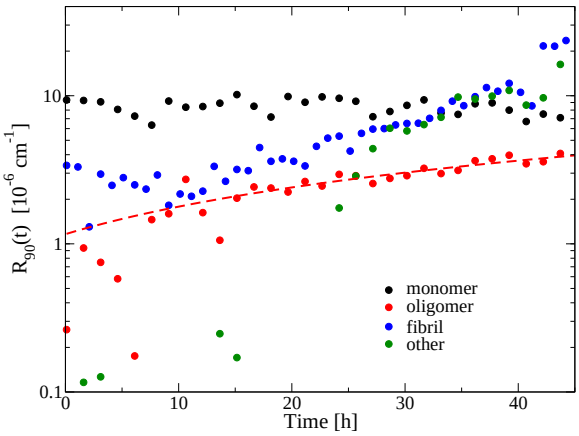

Fig. 4.27: Rayleigh ratio due to species in solution. Incubation temperature 70 $°C$. Dashed line is a guide for the eyes.

is necessary to trigger the formation and the elongation of fibrils. This hypothesis is reasonable because the conformational change of the monomers (on diluted samples) takes more than 2 days.

Further, the oligomers have a net charge larger than monomers. However the surface charge density may be not particularly high [12], and hydrophobic regions are typically clusterized. We expect an enhanced anisotropy in the intermolecular interactions between oligomers, due to the not isotropic distribution of hydrophobic surface patches. Therefore we may argue that oligomers assemble in correspondence of hydrophobic regions trying to minimize the electrostatic repulsion through a stretched morphology. This could be the origin of lysozyme fibrils and could explain why the fibrils need an oligomerization step before growth. It is possible that the oligomerization step is fundamental to enhance the hydrophobic attraction with respect to the electrostatic repulsion. At 70 $°C$, the enhanced speed of reaction (due to the temperature) leads to the formation of amorphous aggregates (Fig. 4.28c) which interfere with the slower growth of prefibrillar aggregates.

4.6 Morphological properties of lysozyme fibrils

We investigated the morphological properties of samples obtained after incubation of lysozyme at acidic pH for one week at different temperatures. We collected some AFM images (Fig. 4.28). They gave us several informations

and were useful to clear better the results obtained by DLS measurements and to try to propose an explanation of the microscopical mechanism that drives the formation of lysozyme fibrils. Fig. 4.28 shows the objects which we find in 18.5 mg/ml lysozyme solutions after incubation and after dilution. In particular, panel a (dilution 100 times) shows a network of fibrils produced by their deposition on mica's plate. Panel b and c show the sample diluted 10000 times. We can see elongated fibrillar structures and sometimes, especially in the samples incubated at 70 $°C$, some globular aggregates. In Fig. 4.28 fibrils seem similar at all temperatures and this is confirmed by next images, at higher resolution. The length of the fibrils varies from few hundreds of

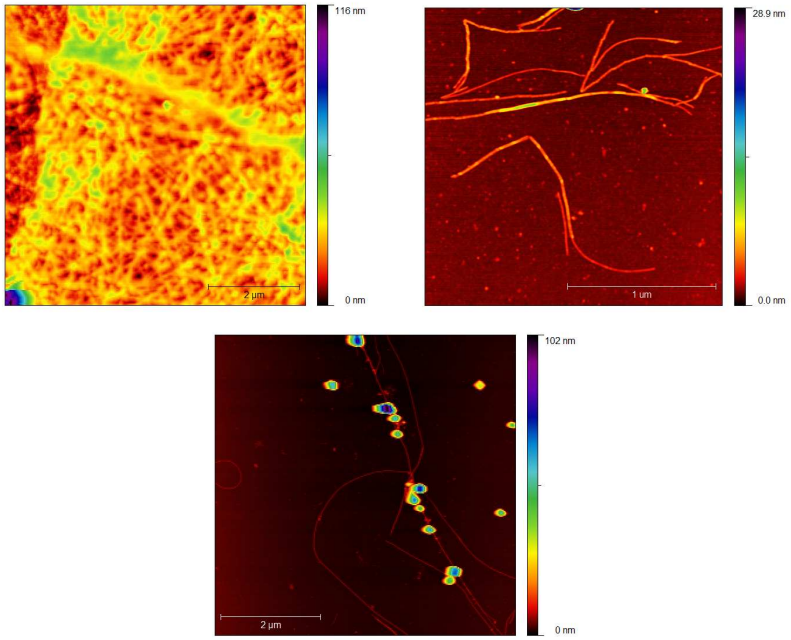

Fig. 4.28: False color AFM images. They show deposited objects on a surface: the rule inside the pictures indicates the length of these objects and colors indicate their height respect to the surface. Lysozyme solution 18.5 mg/ml incubated for 1 week at: a) 60 $°C$ (diluted 100 times), b) 65 $°C$ (diluted 10000 times), c) 70 $°C$ (diluted 10000 times).

nanometers (Fig. 4.29a) to few microns (Fig. 4.28c). The observed diameter of fibrils is quite distributed and we can distinguish simple fibrils and big fibrils due to some thickening process (Fig. 4.29b). The apparent width of the fibrils (orthogonal to the elongation axis) is in the range 25–60 nm and their height is proportional to width and runs from 2 up to 4 nm.

The difference between width and height is not due to some anysotropy of fibrils, but to an instrumental effect. After deconvolution of AFM images, by

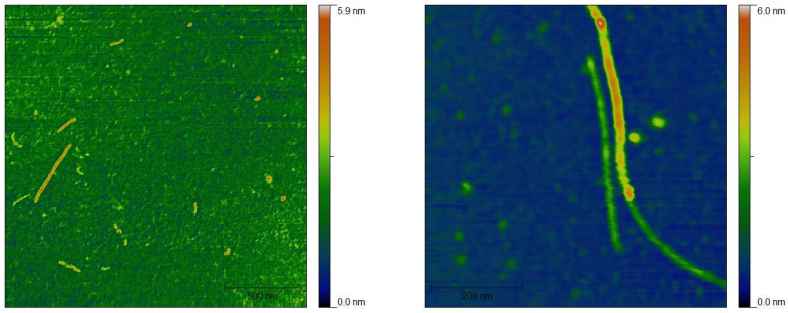

Fig. 4.29: False color AFM images. Fibrils obtained by incubating lysozyme samples at 18.5 mg/ml at acidic pH in the range 60–70 $°C$. Samples were diluted 10000 times before put the sample on the mica's plate. a) Fibrils with a few hundreds of nanometers size. b) Fibrils are characterized by some heterogeneity in thickness.

taking care to consider the geometric properties of the tip model, which we used for each image, we can summarize that there is a certain heterogeneity in the deconvoluted width in the range 20–45 nm. Although our experimental observations are in accord with Arnaudov et al. [7], we disagree about their analysis. It is true that the height is not affected by deconvolution effects of the tip, but the tip-sample interaction is anyway present and affects the height (see § A.9). Thus the obtained numerical value is not a sufficient reason to say that fibrils are formed by simple addition of monomers along their elongation axis. Arnaudov et al. [7] also supported this conclusion because they considered the duration of lag-phase, in scattering experiments, as independent upon concentration. However their experimental data show a slight dependence upon concentration and it seems negligible because they explored a narrow range of concentration (less than one decade). So they excluded a nucleation phase, but this conclusion is not reliable. On the contrary, we give more reliability to thickness of fibrils, estimated by correcting the width (20–45 nm) (as described in § A.9).

If we consider that the oligomer has diameter 20 nm, like the thickness of the thinnest fibrils and guess that their hydrodynamic radius is 10 nm, we can say that the oligomer has a low eccentricity. If we assume an ellissoidal shape (or an equilater cylinder) the other axis is also 20 nm [244]. This is close with the observations of Frare et al. [78], who estimated a spherical shape with diameter 8–17 nm for oligomers of human lysozyme formed at pH 3. From a geometrical point of view, the oligomers with 20 nm in diameter could be made by ≈100 monomers. By analizing the height profile along the elongation axis of several fibrils we observed an axial periodicity around

25–60 *nm*. A periodic length along the axis of the fibrils confirm a twisted or coiled structure. Fig. 4.30, a higher resolution image from Fig. 4.28b, highlights that the fibrils are made of subunits about 120–140 *nm* length and it seems that branching is possible in the junction points because, probably, there are more binding spots in that region than in lateral surface.

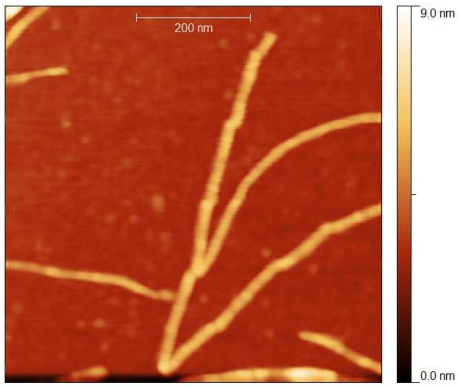

Fig. 4.30: False color AFM images. High resolution image of the subunits in the fibril structure.

Arnaudov et al. have already noted these subunits [7] and observed a coiled structure of the fibrils with a repetition unit of about 30 *nm* by a very sharp tip for high resolution AFM. They hypothesized that the subunits could be due to defects occurring every four or five turns, but there are also other perspectives about that. In principle, they could be an independent brick to build the fibrils or they could be the effect of a very complex and twisted structure made of two or more filaments [261]. We observed these subunits free in solution by AFM (Fig. 4.29a) and the analysis of light scattering correlation functions (Fig. 4.25-4.26-4.27) highlighted a population of objects with a size about 200 *nm* and it becomes significative after two days of incubation. Thus these objects could be the subunits. It is reasonable that if the population of these subunits were sharp and if the biggest fibrils had a size which is a multiple of the subunit size, these subunits could be the building blocks of the fibrils. However these requirements are not satisfied: Fig. 4.29a shows the heterogeneity of short fibrillar objects, which coexist with longer fibrils, after several days of incubation. Thus it is difficult to distinguish experimentally between the idea of subunits as building blocks or a progressive elongation by oligomer addition. I guess that the difference could be in the life time of the populations and could be discerned by AFM imaging in solution, during the incubation. The direct observation of kinetics

of formation of fibril could give a definitive answer about that. We also estimated in a more quantitative way the flexibility of the fibrils.

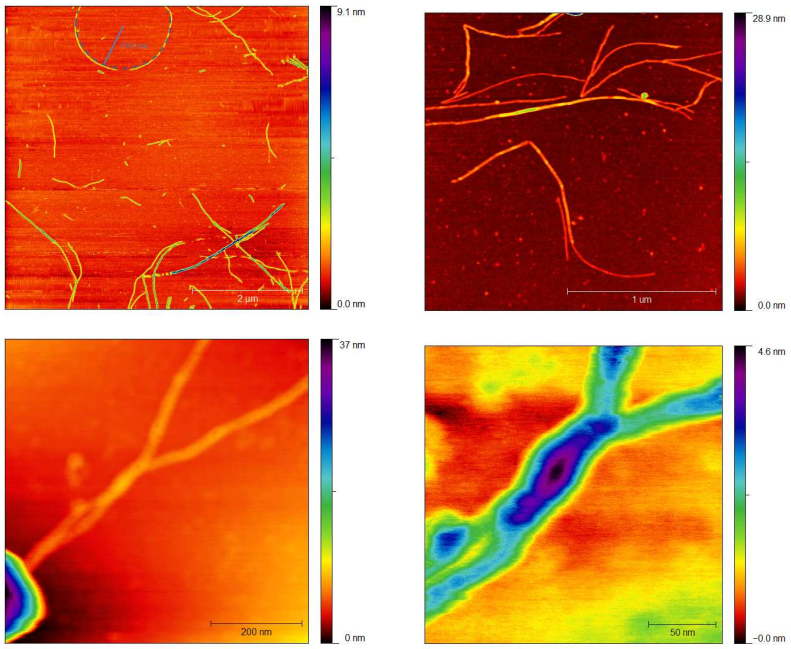

Fig. 4.31: False color AFM images. Panels $a-b$ highlight that fibrils are enough flexible to have a curvature radius between 300 and 900 nm. Panel c shows that flexibility is higher, so great to allow the twisting. Panel d is an high resolution image of a region in Panel c to highlith that it is twisting and not branching.

In Fig. 4.31 we highlighted the curvature radius of some fibrils. Besides we analyzed several AFM images by *2D Single Molecule* software (freeware Nanostructured Materials Group Software. Clarkson University. Potsdam NY; http://people.clarkson.edu/~sminko) to calculate the *persistence length* of the fibrils. It is the length on which the fibril maintains the same direction and it calculated by considering the spatial correlation in the orientation of chains [83]. In my samples the persistence length has a quite high variability because it ranges from some tens of nanometers (as elicited by twisting in Fig. 4.31c) up to a few microns (Fig. 4.31b-4.28c). In particular, Fig. 4.31d highlights the greater thickness in the region of twisting and excludes the possibility that it is a *branching* case.

4.7 Effect of ionic strength on fibrillogenesis

The formation of fibrils is inhibited by the addition of salt in solution and the amount of amorphous aggregates abruptly increases (Fig. 4.32). The reason is that the salt screens the electrostatic repulsions so the aggregation kinetics becomes very fast and do not allow the formation of ordered and organized structures. A recent work by Hill et al. [110] shows that one can obtain lysozyme fibrils at acidic pH at 50 $°C$ by adding some salt. At lower temperature, maybe, conformational change is less and aggregation is not so strong and the right equilibrium between electrostatic repulsions and hydrophobic attractions is restored to build the fibrils.

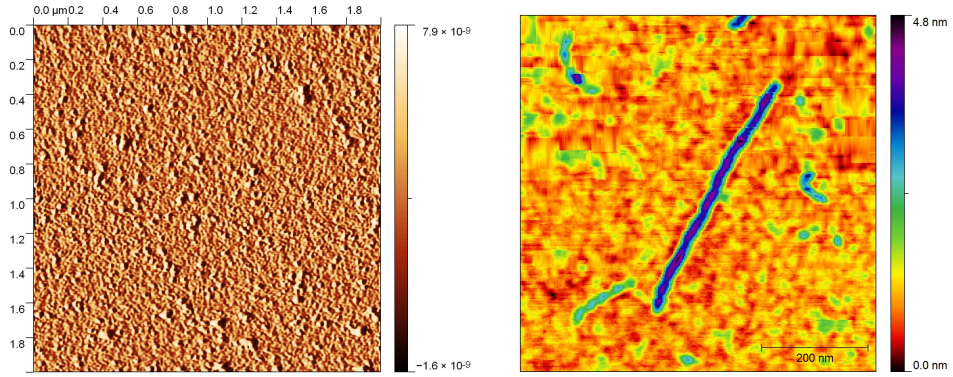

Fig. 4.32: False color AFM images. Lysozyme sample at pH 2 with different amounts of NaCl a) 20 mM; b) 200 mM) and incubated at 65 $°C$. Even if we observed one fibril in the sample with the highest salt concentration, the main population of object is made of amorphous aggregates.

4.8 Conclusive remarks

In this chapter we studied step by step the fibrillogenesis of hen egg-white lysozyme at acidic pH by focusing both on level of single molecule and on intermolecular level. It is reasonable that the instability of a solution, which typically leads to some aggregation pathway, is triggered by instability of single molecules. In our case the origin of this instability is the high net charge of the protein. High charge has two effects: stabilization of solution because of high intermolecular electrostatic repulsion and destabilization of protein because of breaking of surface salt bridges, which are present e.g. at pH 7 (Fig. 4.1a). Also intramolecular repulsive interactions and the resulting deformation of structure concur to destabilize the protein. In these

conditions it is easy to induce a conformational change which further destabilizes the system. In order to study the aggregation process at low pH we increased the temperature. We observed fast conformational changes of secondary and tertiary structure (Fig. 4.1b-c). Structural modifications of the protein change the intermolecular interaction potential. In native conditions, the terms which contribute to the potential are balanced and the solution remains stable. On the contrary, in stress conditions, the delicate balance of these contributions is broken and we addressed our study to understand how this potential changed. In the low concentration regime we found that the second virial coefficient (B_2), an estimate of averaged protein-protein interaction, is independent upon temperature (Fig. 4.5), such as the hydrodynamic effects h_2 (Fig. 4.7b). This is very strange because we expect that at high temperature the hydrophobic effect increases and reduces the repulsive interaction. The behavior of B_2 and h_2 can be quantitatively explained as a first approximation by *Manning* counterion condensation theory. This result is very interesting because it highlights that the very high bare charge of lysozyme at pH 2 is partially screened by counterions, so the effect of pH is almost more relevant on the destabilization of the protein rather than on the stabilization of the solution. Besides, the hydrodynamic radius increases by temperature over 60 $^{\circ}C$ Fig. 4.7a, as already highlighted by Nicoli et al. [171]. According to the electrostatic model which we used, the swelling of lysozyme at high temperature could contribute to increase B_2 and oppose the hydrophobic attraction (Fig. 4.11). We also computed the effect of ionic strength, pH and temperature (in order of importance) on potential generated by lysozyme in a model free framework (Fig. 4.14). Nevertheless very strong repulsive interactions, our system tends to aggregate after long incubation. This observation suggests a competition between repulsive interactions (time independent) and attractive ones (time dependent). The former depend from net charge of protein and the latter depend from its unfolding degree. In this regard, Arnaudov et al. [7] elicited a slow conformational change of secondary structure during some days. We confirmed that (Fig. 4.15-4.16) and we found a similar behaviour at tertiary level (Fig. 4.17–4.21). These slow conformational changes seems to have an important role in the aggregation process. DLS kinetics highlighted that, during the lag-phase, the main aggregation process is the growth of an oligomeric species (Fig. 4.25-4.26-4.27), so oligomers would be the result of competitive interactions. The balance of interactions is so precarious that fibrils are no longer formed at temperature higher than 70 $^{\circ}C$ [7] or in presence of salt (Fig. 4.32) or at pH higher than 3 [7]. In particular we elicited the role of ionic strength on fibrillation pathway. In our system the salt screens the electrostatic repulsion and attractive contributions dominate and drive a disordered aggregation. How-

ever, at low temperature, the addition of salt is necessary to form fibrils because the hydrophobic contribution is not enough [110]. Finally we observed lysozyme fibrils by AFM (Fig. 4.28–4.31). We confirmed some fibril properties highlighted by Arnaudov et al. [7], but we disagree about the analysis on thickness of fibril. In particular our data are compatible with a fibril formed by spherical or cilindric-like oligomers. Other works have evidence of sferical or anular oligomers in lysozyme samples [78]. In the framework of competitive interactions we propose a simple mechanism which could lead the oligomers to assemble into fibrils. Oligomers could be assembled to form hydrophobic ends. In our case, it could be necessary to establish a certain balance between hydrophobic contribution and the electrostatic one. These hydrophobic ends could drive the stacking of the oligomers and the aggregation could go on in an ordered way through the repulsive interaction of the oligomers. We have experimental evidence that fibrillation goes on by addition of oligomers and not by addition of monomers. In fact DLS kinetics at 60 and 65 $^{\circ}C$ (Fig. 4.25-4.26) show that the oligomeric species growths and remains stable more than 3 days by keeping constant its own hydrodynamic radius. If monomers could bind to "oligomers" we would likely observe an oligomeric species with an increasing hydrodynamic radius. Several informations can come from structure of the oligomers. Studies about those are growing, but the question is whether the sample handling can affect the results. Not-invasive techniques are preferred and in the next chapter we will try to reconstruct the $3D$ structure of the oligomer of α-Synuclein mutant.

Chapter 5

Oligomeric precursors of α-Synuclein fibrils

5.1 Introduction

The study of the fibrillogenetic oligomers is abruptly increased in the last years because their deep knowledge of their structure is fundamental to understand the causes and the mechanism of their formation. In this thesis we proposed a plausible mechanism of formation of the lysozyme oligomers at acidic pH. We discussed how a specific balance of physical and chemical parameters can lead to a specific aggregation pathway. We focused on electrostatic interaction and hydrophobic effect and their balance seems to explain quite well the different scenarios we studied. Of course, the number of possible mechanisms which can drive the oligomer formation is large and depends upon the structural details of protein and upon its intrinsic properties.

Because the morphologies ascribed to oligomers are different in literature (elongated for insulin [261], spherical for lysozyme [78], spherical or ring-like for α-synuclein [53]) a more deeply study could highlight some general characteristics in the oligomerization process. We pointed our attention on α-synuclein (α-Sn) because the observations are conflicting and we also want to understand if its oligomerization process has some similar characteristics with the lysozyme one.

5.2 α-Sn, a protein related with Parkinson's desease

In the Chapter 4 we studied some characteristics of lysozyme fibrillogenesis. In this chapter we want to study the structural aspects of fibrillogenesis for another protein, α-Sn. It is involved in Parkinson's desease, but the causes of this pathology are not clear yet [187]. Studies on *post-mortem* brains affected by Parkinson's revealed that this pathology is characterized by citoplasmatic inclusion bodies (*Lewis bodies*) in the compact part of the *substantia nigra*. *Lewis bodies* are mainly made by fibrillar aggregates of α-Sn and thus suggesting a central role of this protein in the desease.

The most accepted idea about the role of fibrillogenetic aggregates is that they are a physiological response of human body to prevent the desease or to make it slower. Indeed scientific community is inclined to think that the toxic species for living cells are oligomeric aggregates, made of a few or a few tens of proteins [185]. There are different hypothesis about the action mechanism of oligomers and they depend from the type of desease. Their toxicity could be due to their small size, so they can enter in the cell and drive it towards apoptosis [37, 185]. In case of anular oligomers, they could create pores in the cell membrane and affect ionic equilibrium between the inner and the outer part of the cell [136]. This hypothesis has been also formulated for β-amyloid, poly-glutamine and lysozyme [117]. Fibrils could be a smart natural way to get out oligomers by packing them in bigger aggregated structures [263].

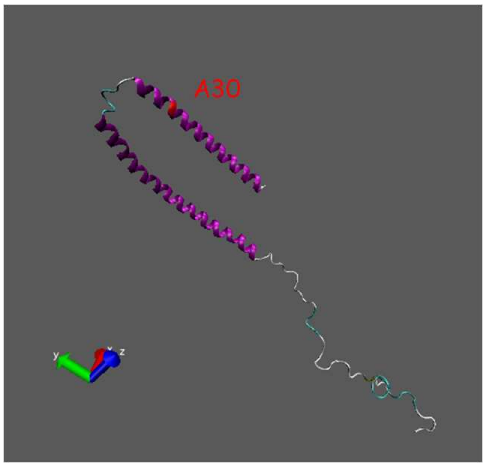

Fig. 5.1: NMR structure of human α-Sn when it is bound to lipidic micelle. From Protein Data Bank 1XQ8

α-Sn is present in the membrane of the neurotransmitter vescicles, which are used to release the neurotransmitters in the synapse. It seems that α-Sn modulate the organization of lipidic components in the membrane of vescicles and inhibit the lipid oxidation [254]. Also, it can act as a molecular chaperone of synaptic proteins. All these biological functions and the possibility to bind several proteins and ions are justified by its intrinsecally unfolded structure (Fig. 5.1) [254, 217]. α-Sn has 140 residues and weights 14.5 kDa, likely lysozyme. It contains three regions [254]:

- N-terminal region (1–60): it is amphipatic and contains four 11-residues imperfect repeats;
- central region (61–95): contains the highly aggregation-prone NAC (Non-amyloid β component) sequence, found abundant in Alzheimer's plaques;
- C-terminal region (96–140): it is rich in acidic residues and prolines.

The region 6–37 has a preference for helicoidal conformation and it takes this structure mainly when it is bound to the lipids of vescicles [254].

We know that amyloidogenic pathologies are often related with puntual mutations of some proteins. The most known pathogenic mutants for α-Sn are A53T (alanine 53 substituted by threonine), isolated from an italian family and A30P (alanine 30 substituted by a proline), isolated from a german family. In particular we studied the fibrillation process for mutant α-Sn A30P. Alanine and proline are both apolar, but the proline destabilizes the protein structure because it is a cyclic amino acid and cannot be connected to other amino acids by a peptidic bond, but only by an hydrogen bond. For this reason proline cannot stay in a α-helix or in a β-sheet and often it is located in turns or in that regions which need more mobility. Mutation A30P decreases the propension for α-helix and slows the rate of fibril formation.

5.2.1 Role of mutations in the aggregation process

ThT assays were performed on incubated sample of wild type α-Sn [89]. The protein solution was 12 mg/ml in phosphate buffer 20 mM at pH 7.4 with 150 mM of added $NaCl$. The sample was incubated at 37 $^{\circ}C$ with ThT 40 μM). The same experiment was repeated on the A30P α-Sn. Fluorescence plate reader has been used to monitor the ThT binding kinetics. A comparison of ThT kinetics between wild-type and mutant α-Sn shows that the wild-type kinetics reaches a bigger plateau than mutant (Fig. 5.2). If we scale suitably the kinetics of the mutant we can observe that both kinetics

are superimposed and this means that the ThT binding process is the same in two both systems. In particular they have the same lag-phase and the same growing rate. However the amount of bound ThT is lower in the mutant and this can depend from a lower total surface of fibrils exposed to the solvent or from a lower ThT binding efficiency.

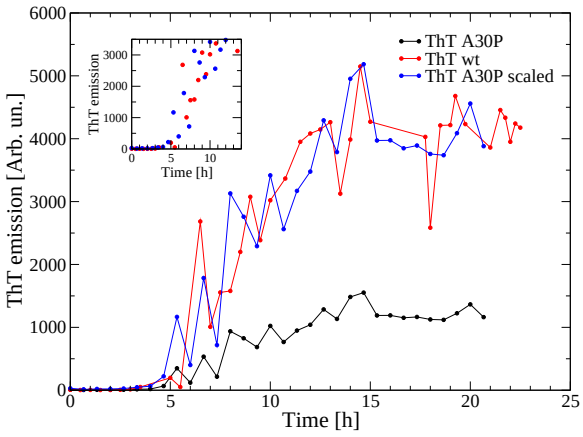

Fig. 5.2: Comparison between binding kinetics of ThT (c = 40 μM) in wild-type and mutant α-Sn sample. Inset is a zoom. ($\lambda_{exc} = 450\ nm$; $\lambda_{em} = 485\ nm$)

This result confirms the previous results by other groups, who concluded that fibrillogenesis of α-Sn A30P is inhibited with respect to wild type. This may be due to the stabilization of oligomers with respect to fibrils [39]: then oligomer may be less stable than fibril, the energy barrier from native to oligomer could be lower than one from native to fibril. It is worth to recall that fibril deposit within the brain of asymptomatic patients is about ten fold higher than deposit in exhibiting patients . That encourages the current hypothesis that toxicity is mainly due to oligomers, while fibrils have only a secondary and less pathological role. A similar effect on fibrillation kinetics (lower plateau) is found if baicaleine is added to wild type α-Sn samples [114]. This compound stabilizes the oligomer, accelerates its formation and slows the fibrillogenesis. If this interpretation were correct, baicaleine could accelerate the disease. The recently suggested fibrillogenesis mechanism is nucleated polimerization (*Oosawa-Asakura* model), able to explain some relevant characteristics: no fibril formation below a critical protein concentration, lag-phase (it can be deleted by seeding), strong dependence of the growth rate upon the protein concentration [190]. The nucleus is kinetically accumulated with respect to monomer above the critical concentration and polimerization of oligomers into fibrils can run [190]. This is the reason for

the chosen concentration in our experiments. This model does not work for protein concentrations higher than supercritical value. In this case the nucleus is more energetically favoured than monomer and the fibrillogenesis rate is concentration-independent [190, 5]. According to the most accepted idea, the nucleus is the building block for elongation [261, 135]. Therefore it is very important to know its structure. A great help in investigations comes from direct observation techniques. Wild type and muted α-Sn samples were studied and observations about structure of oligomers do not match each other. According to Lashuel et al. [137] A30P mutant forms anular protofibrils, A53T forms anular and tubular protofibrils and wild type forms anular protofibrils, after a long incubation. According to others [47] A30P mutant forms spherical oligomers at the beginning, while at long times a different pathway is suggested: formation of short chains and finally anular structures [48, 53]. They use similar conditions than ours (pH 7.4, about 100 mM $NaCl$), but a bit lower concentrations. It is not clear yet if these oligomer morphologies are consequential or if there is some morphological bifurcation in the pathway of oligomer formation. So, it seems that some heterogeneity in oligomers is revealed [137]. That heterogeneity is not only about oligomer structure, but also about its size. For example anular structures with diameter in range 8–24 nm for A30P were observed [144].

5.3 Isolation of oligomers: the method

During a fibrillation kinetics the species in solution are different in size and morphology, but tipically the main components are distinguishable only in an indirect way because the equilibrium is dynamic. The current idea, strongly supported by experiments, is that monomers, oligomers and different sizes of fibrils are present in solution. Each of these species is characterized by a scattering pattern and if we have accessory informations, like form factor at the beginning (*monomers*) and in the late stages of kinetics (*fibrils*), we could isolate the form factor of the *oligomer*. This is the key point of our method and the most complex step. The good reliability of this method, of course, depends upon the heterogeneity degree of the sample. After we find the form factor for the oligomer [126], it is quite simple to model the $3D$ structure of the oligomer. In fact a software could compare the form factor produced by a bead-model with the form factor from the oligomer [240]. The procedure can be quite long, but it can be automatized. The result is not unique, so the last critical step will be to average the founded models by clustering the most similar ones and choose the most reliable [262]. The analysis is not univoque because of the noise of data and some heterogeneity

in the sample, but crossed controls can be very useful to discard the wrong models and have a good probability that found solution is correct. This approach was used successfully in the determination of structural nucleus of insulin amyloid fibrils [261].

5.3.1 Behaviour of the solvent

The most important step to perform a good analysis is to subtract the buffer signal from the sample signal. Data were collected by alternating the sample and the buffer. So we have the buffer signal before and after each measurement. This is important for a correct evaluation of data because the intensity of X-ray beam is not constant in time. So the best way to do that is to average the two both buffer measurements.

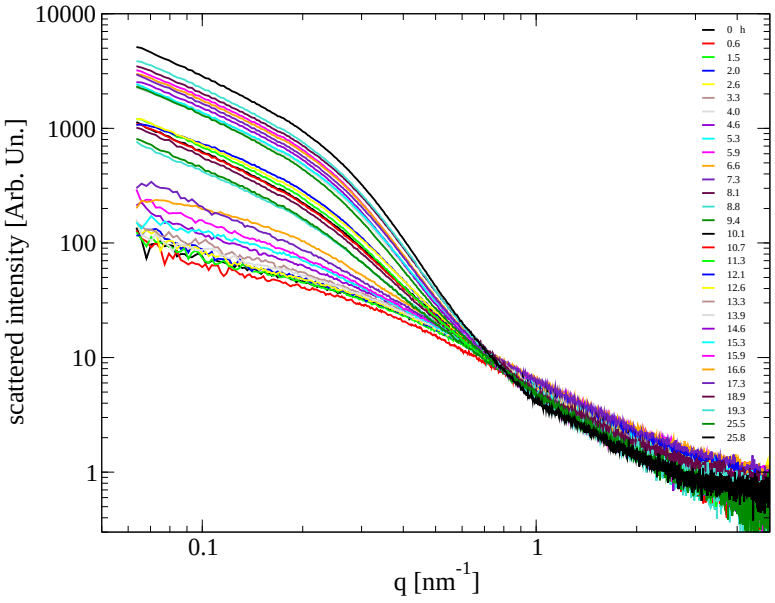

Fig. 5.3: Kinetics of SAXS buffer subtracted form factors: scattered intensity vs modulus of wave vector.

SAXS form factors, buffer subtracted, show an increase of diffraction pattern at low q-values. At high q-values structure functions become lower and lower (Fig. 5.3). In this q-range the origin of this effect is surely due to some modifications of the buffer. The only modifications that a buffer can undergo are in its structure around proteins or aggregates. In particular we tried to correct the buffer-subtracted data $F_b(q)$ for a linear function:

$$F' = aF_b(q) + b \tag{5.1}$$

where a is a multiplier to correct the concentration (eventually modified by some precipitation effect) and b is a constant that can correct a different arrangement of solvent around aggregates (Eq. 5.1). The basic idea behind this correction is that at high q-values the signal of the protein cannot be modified so we expect that all the structure functions are superimposed in that q-range. We guess also that a fibril changes the properties of the local environment because of exposure of hydrophobic regions, used to bind other building blocks. So the water molecules closer to the aggregates could organize in an ice-like structure, just to bury the hydrophobic regions from the rest of the solvent. We expect some trends in a and b-values during kinetics. We guess that these corrections are small at the beginning of the kinetics and grow in time. We chose the smallest corrections as possible among the different possibilities. We obtained a regular growing trend for a-values, related with concentration. The b-values were quite random. The reason for that maybe is that the relative variations on the buffer signal are small, so the buffer subtraction operation hides these features. Therefore we tried to perform again this correction on the raw data, not subtracted by buffer in the region 2.34–4.5 of q-values. We tested four corrections by imposing the first values of a or b; in fact the next values in the kinetics are quite constrained, if we respect the principle "*small corrections as possible*". In particular we tested the overall range of $a - b$ values that can correct succesfully the data: $a = 1$, $b = 0$, $b > 0$, b more negative than case $a = 1$.

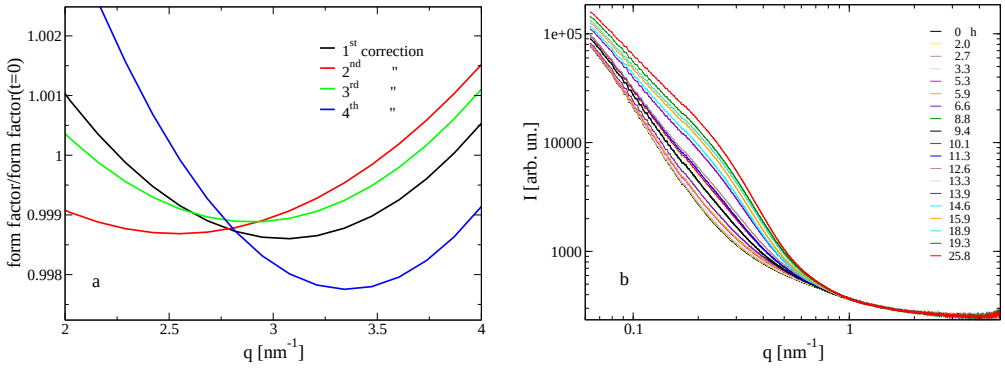

Fig. 5.4: a) q-dependence of corrections; b) raw form factors corrected by correction 1 according to Eq. 5.1 (with a(0)>1 and b(0)=0).

$a > 1$ means that there is a bit of precipitation so we are guessing that the real protein concentration in solution is lower than initially measured. We considered also the case in which $a < 1$ just to take into consideration of possible errors in the determination of the initial concentration. We observed a progressive decrease of form factor signal at high q-values and if we guess

that this decrease is due to buffer reorganization, we must assume that $b > 0$, that is water is ordering around fibrils. $b < 0$ has no physical meaning in this context, because water is in the "liquid form" at the beginning and it cannot be more "liquid" than it is. Negative values for b could only take into account some shifts (blank, etc.). We tried, just as a test, to plot the ratio between corrected form factors (with respect to the first one) and the first data set (after two hours incubation) in the q-range 2.34–4.5 (Fig. 5.4a). First ($a(0) > 1$ and $b(0) = 0$) and third ($a(0) = 1$ and $b(0) < 0$) correction seem to have less variations in the considered q-range and we choose the first correction for qualitative reasons. Indeed a-value starts from a bit more than 1 (because a bit precipitation could be after 2 hours) and grows; b-value starts from 0 (because initially water is liquid) and progressively grows because of the ordering of water around fibrils. Then we corrected the raw data (Fig. 5.4b) and subtracted the buffer signal from each data set and we normalized by protein concentration (Fig. 5.5). In this way we obtained a data set comparable (except for concentration normalization) to Fig. 5.3, buffer subtracted but without corrections.

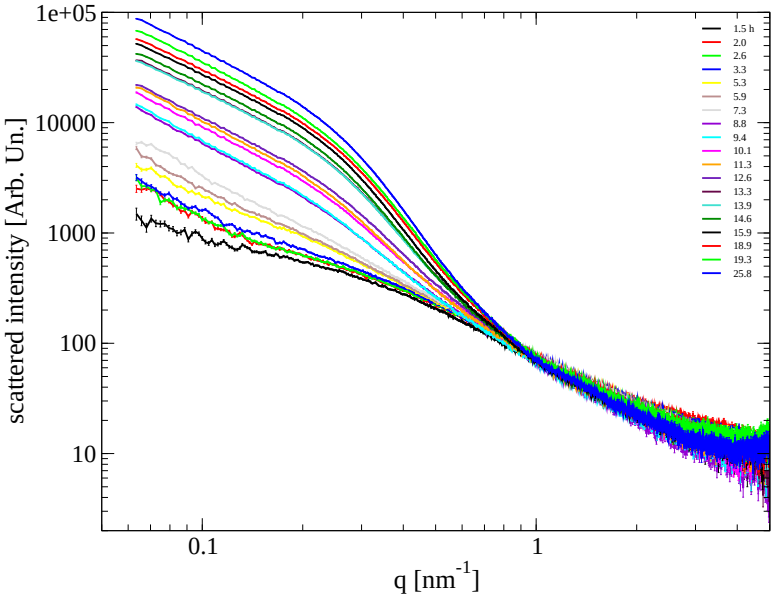

Fig. 5.5: Corrected, buffer subtracted and concentration normalized form factors.

We analyzed these corrected data by *PRIMUS* [126] a software included in *ATSAS* [125], a software suite developed in the Svergun's lab. We extracted some important parameters like R_G , I_0, D_{max}, respectively the *Guinier* radius, scattered intensity at zero angle and maximum diameter of objects

in solution. R_G, I_0 and D_{max} increase by time as expected. It is worth to note some quantitative differences in R_G and I_0 before and after correction, but the qualitative behavior remains the same.

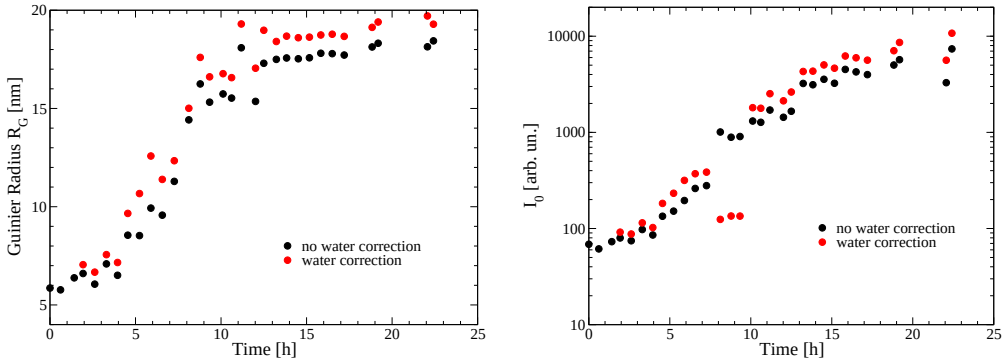

Fig. 5.6: R_G and I_0 behavior with or without water correction of data.

If we scale suitably D_{max} kinetics by a sigmoidal model, we can note that ThT, D_{max} and R_G have the same behavior. R_G gives an average information about size of species in solution, but D_{max} is strongly indicative of the size of aggregates. The correlation between ThT and D_{max} is in accord with the quite well accepted hypothesis that ThT binds to the biggest species in solution, because it needs to link to specific sites created by macromolecules. However the correlation between ThT and R_G could be misleading.

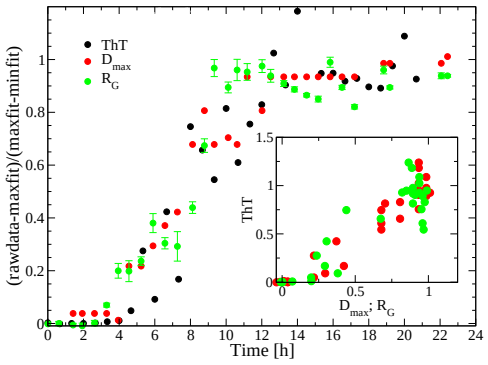

Fig. 5.7: Comparison among ThT emission kinetics, D_{max} and R_G. Inset: correlation between ThT signal and D_{max} (red) or R_G (green)

Then we calculated the Fourier transform of corrected data by $GNOME$, another software package included in $ATSAS$ [125], and we obtained the pair distribution function $P(r)$, the probability density distribution to find two

atoms at distance r. We observe that area under these functions increases by time, because the area is proportional to the square volume of the objects (Eq. A.10). We normalized these distributions by their own areas to better visualize the features of all curves and to give a statistical meaning to these functions. In fact, these normalized curves are the density probability. The area under the curve between R and $R+dR$ indicates the probability to find "distances" between R and $R+dR$ inside the objects in solution.

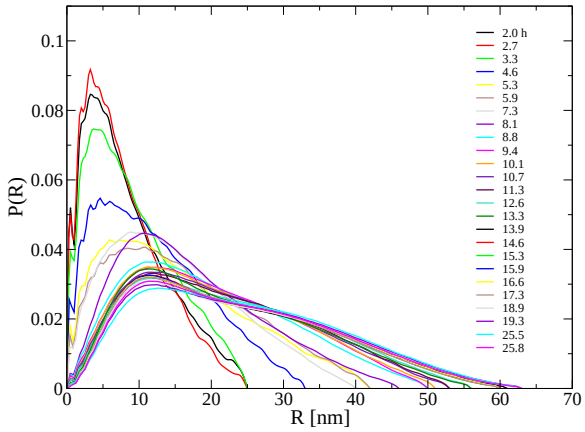

Fig. 5.8: Volume normalized pair distance distribution, obtained by Fourier transform of data in Fig. 5.5

Besides this way to represent pair distance distribution function $P(R)$ can highlight differences in the form of the objects. It can say if the objects increase their size without changing their form, or if there is some morphological transformation. In our data interpretation of $P(R)$ is quite difficult because our sample is not monodispersed, but there are different sizes objects. So modifications in $P(R)$ are related with changes in volume fractions of species in solution.

If we compare SAXS curves (Fig. 5.5) to ThT kinetics (Fig. 5.2), we highlight that at the beginning (before 4 h) and at the end of kinetics (after 15 h) the ThT signal is flat (within experimental errors), even if SAXS curves change a bit during these phases. In particular, we trust that the lag-phase is characterized by structural changes of monomers or oligomers, which are insensitive to ThT binding because they are not enough big to have the suitable binding sites. The final plateau could be characterized by some repacking of fibrils.

5.3.2 Multi-component analysis

Our goal is to isolate the form factor of the oligomer, so we need to know how many components are present in solution. We made that by *Singular Value Decomposition* (SVD) algorithm. It is a mathematical procedure used to factorize a matrix [209]. Suppose that A is an $m \times n$ matrix whose entries come from the field $\mathbb{C}$. Its columns could be considered a basis of vectors: in our case they can be the experimental form factors. It is possible to find an $m \times n$ diagonal matrix S with p non-negative real numbers ($p = min(m, n)$) on the diagonal such that

$$A = USV^{\dagger} \tag{5.2}$$

U and V are respectively $m \times m$ and $n \times n$ unitary matrices ($U^{\dagger} = U^{-1}$, $V^{\dagger} = V^{-1}$) and $V^{\dagger}$ is the hermitian conjugated matrix of V. The values s_{ii} on the diagonal of S are called *singular values* of A and typically are ordered in a descending order. The columns of U are the normalized eigenvectors of $AA^{\dagger}$ and represent a new basis of vectors. In our case they represent the mathematical components which we could use to decompose linearly the experimental form factors. However these components have not a physical meaning but suggest the number of components which are necessary to reproduce data within a certain accuracy. The columns of V are the normalized eigenvectors of $A^{\dagger}A$. The algorithm gives also the relative weight that the new components have in the linear combination. We used the program $SVDPLOT$ [126]. If we consider all curves this analysis suggests that we need four components to decompose data. But the fourth is close to the irrelevant ones. For this reason we tried to "clean" the data sets by hope to reduce the fourth component. We did not consider some data sets because they are quite strange respect to trend followed by others. Then we tried to delete high q-values region because it is noisy and the signal is low and the SVD plot does not depend signifivatevely from this. So we decided to delete as less points as possible ($q > 3$). We also tried to delete some first points because they can be affected a bit by incident beam (*stray light*). By looking at eigenvalues, we report that the second component is always the same. If we delete too much points all components are reduced. If we delete only first or first two points at low q, the fourth component is strongly reduced. In particular, if we delete first two points the fourth component is random. If we use these "simplifications" and perform again SVD analysis by all data sets, eigenvalues plot indicates three main components, but eigenvectors plot shows that the fourth component is not random (Fig. 5.9). By summarizing, the "set-up" to continue the analysis is the following: 1) some "strange" files eliminated; 2) first two points at low q are deleted; 3) points beyond ($q = 3$) are deleted. Finally we also tried to test what happens if we do not consider

the last data sets in the SVD analysis. The basic idea behind this test is to try to understand if there is some repacking of fibrils. We were helped from ThT kinetics to do that and we selected the data file corresponding to the beginning of the final plateau (after 16 h). If we do not consider these final data sets, the SVD-plots remain qualitatively the same. If we neglect some more data sets, in correspondence of the elongation phase, the third component in SVD analysis is strongly reduced and only two main components are evident. So this analysis allows us to conclude that there are three main components and there is no repacking of fibrils or it is not so strong. We guess that the three components are monomers, oligomers and fibrils. It is

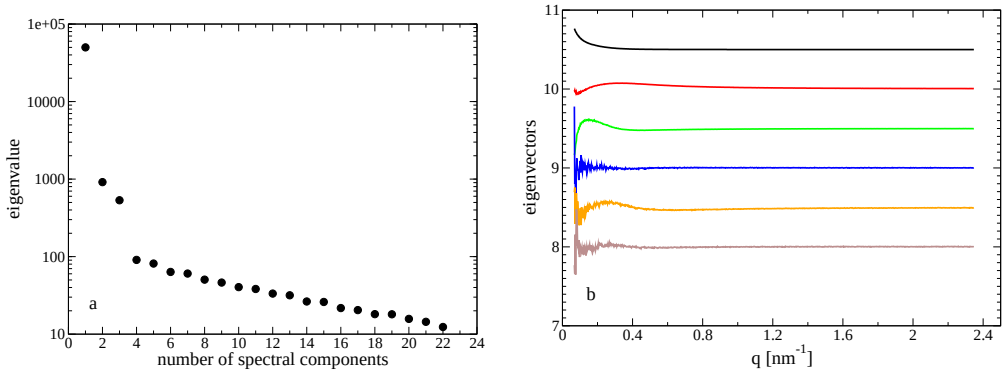

Fig. 5.9: Single value decomposition of form factors. a) shows the *eigenvalues*; b) shows the most relevant spectral components (*eigenvectors*).

worth to consider just a warning in the SVD analysis. It does not consider error bars in the data sets and we have a lot of other components quite at the same level of reliability. So, it could be necessary to use four components in the decomposition.

5.3.3 Decomposition of data

Proofs with three components

If we know the form factor of monomer and fibril, we could find the form factor of the oligomer. α-Sn is a destructured protein: wild-type and mutant are different in one residue, so we do not expect relevant differences in the form factor of the wild-type or the mutant monomer. We do not have the experimental form factor of the mutant, so we will use wild-type monomer for the decomposition. We can use the wild-type oligomer as a starting point to model the mutant oligomer. However we need the right form factor of fibrils. We used the program $OLIGOMER$ [126] and tried to use each

form factor of fibrils correspondent to final plateau in ThT kinetics. The best decomposition attempt is obtained if we use the last form factor of the kinetics. Unfortunately the fitting procedure does not converge: the residuals of fits do not have the same features for each curve. Our goal was that the form factor of oligomer could converge to a single function. If we look at volume fractions obtained, we can highlight the curves in which the fraction of oligomer is higher. These curves, of course, have a higher reliability to extract the form factor of the oligomers. However, even if we consider only the form factors with an high volume fraction of oligomers, the convergence is not guaranted. Fig. 5.10 shows that the monomer fraction decreases, the oligomer fraction decreases after an initial plateau and fibril fraction increases. In particular we can note that ThT has the same behavior of the volume fraction of the fibrils and this strongly confirms the current idea that ThT is preferentially bound to fibrils (Fig. 5.10) [165, 139, 95, 264].

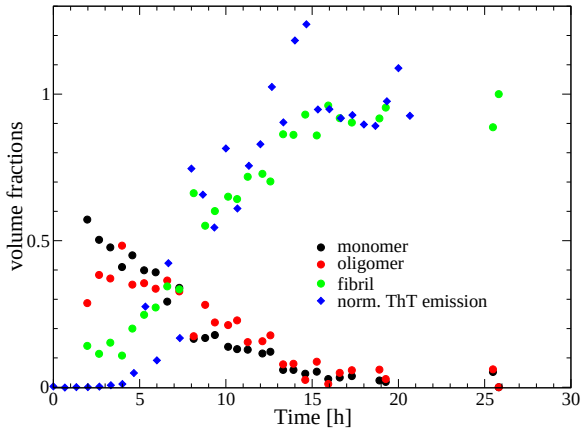

Fig. 5.10: Estimated volume fractions of monomers, oligomers and fibrils as obtained by a three-components decomposition.

Therefore, regarding to Fig. 5.7, we can confirm that the real physical information is into the correlation between ThT signal and D_{max} trend; the correlation between ThT and R_G, which could suggest the binding of ThT to smaller species, is due to the average on all species in solution (the biggest ones give an higher contribution).

We tried also to fit data without using wild-type oligomer as an initial guess or using the first form factor collected during mutant kinetics, but the convergence is not better and showed that the results do not qualitatively change .

Proofs with four components

The best approximation for the A30P α-Sn oligomer, obtained by decomposition through three components, is characterized by a steep increase at low q values, so fibril form factor, probably, is not correctly considered. The best hypothesis is that there is some heterogeneity in the sample. SVD analysis allowed us to conclude that, probably, the repacking of fibrils is not so strong at first approximation. In conclusion, three components are not sufficient to describe the experimental form factors. Therefore, we tried to consider two types of fibrils. We fit the data by using wild type monomer, wild type oligomer as initial guess, the mutant starting form factor and each one among all others data curves (one by time) but the convergence was not so better than decomposition attempts by three components. We also tried to use wild type monomer, wild tipe oligomer as initial guess and several combinations of the two form factors from the final part of kinetics. The best attempt is to start with:

- the form factor of the wild type
- a form factor among the first in the kinetics of mutant
- the form factors collected at the onset and at the end of the kinetics of the mutant.

Although we added one more component, the quality of the results was not improved. It seems that there are not three or four well-defined species, but there is an intrinsic heterogeneity, difficult to be considered in a suitable way by analysis procedure.

5.4 Conclusive remarks

In this chapter we studied a kinetics of form factors about fibrillogenesis of α-Sn A30P. By comparing kinetics of ThT fluorescence emission, we confirmed that the fibrillogenesis of the mutant is inhibited with respect to the wild-type (Fig. 5.2). Probably this is due to a greater stabilization of oligomers with respect to fibrils and could explain the high toxicity of the mutant α-Sn. We studied the behavior of water around fibrillar species and we found that some water molecules tend to organize in ice-like structures around the fibrils (Fig. 5.4). Although SVD analysis suggests that three components are sufficient (Fig. 5.9), this is true only as first approximation. By this warning, we decomposed the experimental form factors by using the form factor of monomer, oligomer and fibril and we obtained a confirm that ThT binds

to fibrils, not to prefibrillar species (Fig. 5.10). However, strictly speaking, the decomposition of SAXS data was not achieveable with a reduced set of form factors. The existence of a low amount of fibrils at the beginning and repacking at the end (heterogeneity of fibrils) are possible sources of smearing. However the residuals of fit have not always the same features during kinetics and this suggests that some dynamic equilibrium of different species (more than three or four) may be present in solution. We did not take into account the possibility of some heterogeneity in oligomers too. The hypothesis, experimentally supported, by Samori's group is that the structure of α-Sn monomer fluctuates among three different conformers, characterized by different amount of β-sheet and *random coil* [222, 27]. In particular the mutants have an higher β-sheet content. According to them, the experimental conditions and the type of mutation modify the equilibrium among these conformers and drive proteins in a particular morphology of fibrillation. Thus, even if we have always different types of oligomers, the highest population depends upon the initial conditions. Some oligomers lead to the formation of elongated fibrils. Others oligomers are stabilized in pore-like structures [53]. If there is a constant equilibrium between different oligomeric structures we would find an "average" structure, but if the equilibrium changes kinetically during the time only one component is not sufficient to describe the oligomer structure. In this case we should have structure function about one of the oligomers, so further investigations about heterogeneity of sample and methods to stabilize only one oligomer type are necessary.

Chapter 6

Conclusions

Except some cases where proteins form fisiologically aggregated structures, like collagen fibers, protein solutions are stable in physiological conditions. Proteins form aggregated structures because of some physical and chemical changes of the environment or mutations (e.g. HbS). All these factors affect the structure of the protein and destabilize this and the solution. Several interactions are involved at intra- and intermolecular level. Some of them are attractive (van der Waals, hydrophobic), others are repulsive or can have an ambivalent character. We focused on the balance between the repulsive electrostatic contribution (at long range) and the attractive contribution due to exposure of some hydrophobic regions of the protein. The attractive contribution is typically modulated by conformational changes of the protein (e.g. caused by temperature, pH, mutations, etc) or by cosolvents added into the solution (e.g. TFE). The repulsive contribution is modulated by pH and ionic strength.

The aim of our thesis is to find new insights about the role of a different balance among the intermolecular interactions of proteins on different aggregation pathways. We led our investigations by studying, mainly, lysozyme solutions. Only the last chapter is focused on mutant α-Synuclein to try to make clear some aspects.

In the Chapter 2 we studied hen egg-white lysozyme at physiological conditions eliciting the role of different physical and chemical parameters: temperature, ionic strength, protein concentration and added alcohol. In the first section we studied the role of heat around the unfolding temperature. We found that the unfolding transition, revealed by DSC measurements, is affected by the aggregation (Fig. 2.5). Besides, we demonstrated that the aggregation process leads kinetically to the thermodynamic irreversibility (Fig. 2.2-2.10). We analyzed more deeply, through DLS data, an aggregation kinetics at temperature where the unfolding degree is mild and revealed the

increase of size of aggregates and their temporal dependence (Fig. 2.8). We found a so strong correlation between the enthalpy change of the unfolding transition and the energy barrier of aggregation that we can argue that unfolding is the rate limiting step of aggregation. Then, we studied the unfolding transition in conditions where the aggregation is strongly inhibited. In order to do that, we performed the same set of experiments at very low ionic strength and we highlighted that unfolding is reversible.

In the second section we studied the role of protein concentration upon the aggregation mechanism of thermally denatured proteins. We found that basically the mechanism is the same through turbidity measurements (Fig. 2.16). Structural analysis of aggregates, through SALS measurements, highlighted the existence of small compact aggregates about few microns at all concentrations. They are organized in superstructures with size dependent by protein concentration (Fig. 2.18). In particular rheological measurements allowed us to observe clearly a threshold beyond which the system gelificates (Fig. 2.22). Viscosity measurements revealed some differences dependent upon concentration, but at infinite shear rate the value is the same (Fig. 2.21-2.22): this observation suggests that large clusters are destroyed by mechanical stress because connecting interactions are weak; on the contrary, the small compact aggregates remain at all concentrations.

In the Chapter 3 we studied the role of alcohols on the stability of lysozyme. We observed the preferential hydration of the protein through the observation of hydration of Trp 62 (Fig. 3.1b). Far UV CD spectra and fluorescence spectra highlighted that increasing amounts of TFE do not change the secondary and tertiary structure of lysozyme at room temperature (Fig. 3.1a). However the increase of temperature changes secondary and tertiary structures. At low concentration of TFE, CD spectra show an isodichroic point as a function of the temperature (at 204 nm), a footprint of a two-states transition (Fig. 3.3). However the lost of sharpness in the isodichroic point and the very strange spectrum at 20% of TFE suggests the presence of an intermediate state of unfolding. This is confirmed by the more consistent analysis performed by a three-states model which allows to fit together all the sigmoidal curves (Fig. 3.5). We can also confirm the typical stabilizing effect of TFE on the secondary structure of lysozyme (Fig. 3.6). The incubation of samples with different amounts of TFE at high temperature highlights the presence of fibril-reach solutions, as shown by ThT assays or direct visualization of samples (Fig. 3.7-3.8). The presence of fibrils was also confirmed by AFM imaging, notwithstanding the difficulty to extract some single fibrils from the bundles which, maybe, are extended over the whole sample (Fig. 3.9).

In the Chapter 4 we studied the fibrillogenesis of lysozyme at pH 2. A fast

conformational change of secondary and tertiary structure occurs by increasing the temperature (Fig. 4.1b-c). We observed that lysozyme is destabilized at pH 2 with respect to pH 7, because of a different distribution of electric charges (Fig. 4.1a). In order to study the intermolecular potential, we measured isothermal compressibility as a function of the temperature through static light scattering in the low concentration range (Fig. 4.2). We also measured the collective diffusion coefficient of lysozyme in our system (Fig. 4.3). In particular, we found that B_2 remains constant (Fig. 4.5) and the hydrodynamic radius increases (Fig. 4.7a) by temperature. We obtained an estimate of hydrodynamic coefficient, which takes into account the long range viscous effects (Fig. 4.7b). We explained the values of B_2 by using mainly an electrostatic contribution which has the Yukawa form, but an effective net charge of the protein calculated through the Manning condensation theory. However some aspects suggest that the electrostatic contribution is not the only one, although it is dominant. In fact the simulated values of B_2 show a decreasing trend by the temperature (Fig. 4.11). We performed some simulations by using a Non Linear Poisson-Boltzmann equation solver to try to visualize some aspects in an eventual anisotropy of surface density charge (Fig. 4.13). By performing a spherically average of the obtained potential, we elicited that ionic strength, pH and temperature are important in a decreasing order (Fig. 4.14).

We focused on kinetic conformational changes observed during lag-phase of the fibrillogenesis. In particular we observed the exposure of tryptophans and a related change in the tertiary structure in the temperature range 60–70 $^{\circ}C$ (Fig. 4.17-4.18-4.19). Similarly we observed kinetic changes in the secondary structure by CD (Fig. 4.15). Besides aggregation kinetics (at high concentration) highlight the formation of oligomers with a size about ten nanometers (Fig. 4.25-4.26-4.27). They are formed early, during the lag-phase through a saturation exponential behavior. Their concentration seems to stay in equilibrium for several hours. After about two days we can see the formation of fibrils. It is not clear yet if fibrils are present before two days (but we cannot see them because of the small number) or if they starts really to form after two days. Interesting hypothesis can be formulated dependently from that. Finally we observed the structure of fibrils after several days of incubation at high temperature in the range 60–70 $^{\circ}C$ (Fig.4.28–4.31).

In the Chapter 5 we analyzed some SAXS form factors about the fibril formation of the mutant α-Synuclein A30P. First, we compared ThT kinetics between mutant and wild type and confirmed that the mutant has an inhibited kinetics (Fig. 5.2). We started the analysis of SAXS since buffer correction). It is not a simple subtraction operation and this procedure allowed to highlight the change of solvent structure by the time (compare Fig. 5.3

and Fig. 5.5). We found the minimum number of components, necessary to describe completely all the curves: they are three and we assigned them respectively to monomers, oligomers and early fibrils (Fig. 5.9). Therefore if we have the experimental form factor for monomer and fibrils, we could obtain, in principle, the form factor of the oligomer. The convergence of the form factor of the oligomer is very critical. We tried to use four components and a lot of combinations, but the result is not improved. The best fit obtained (by three components) provides for an estimate of the volume fractions of the components (Fig. 5.10). We conclude that oligomers and fibrils exhibit an intrinsic structural heterogeneity.

In conclusion, our study revealed a strict connection between the physical-chemical parameters of the protein system and the aggregation pathways by highlighting that the key-point is not the individual interactions, but their balance. In fact, different conditions exist where a protein has the same (or similar) behavior, although the single contributions are very different. This concept allowed us to propose an explanation for lysozyme fibrillogenesis at acidic pH. In addition to general importance, it has a practical utility because it can drive people, in a conscious way, to study the same process in different conditions.

Appendix A

Experimental procedures

A.1 Sample preparation

Hen Egg-White Lysozyme (three times cristallyzed, dialyzed and freeze-dried) was purchased from Sigma Chemical Co. and used without further purification. We used three types of solvent at pH 7: buffered solution (with or without TFE) and water corrected at pH 7. The buffer solution was 0.1 M Na-phosphate in *Millipore Super-Q* water. The samples in presence of TFE were prepared by adding suitable amount of TFE, preliminarly dissolved in the same buffer solution to avoid changes in the ionic strength of buffer. Protein powder was dissolved into buffer. Sample was centrifuged at 800 g for 2 minutes to remove air bubbles and small amounts of undissolved material and filtered through a 0.2 μm Sartorius syringe filter. Water solutions at pH 7 and solutions at pH 2 were prepared by dissolving the protein powder in *Millipore Super-Q* water and pH was corrected by adding small amounts of 1 M hydrochloride acid. Sample was filtered into a cell through a 0.2 μm *Millex LG* syringe filter. Sample concentration was determined by UV absorption spectroscopy measurements (Shimadsu UV-2401 PC). The extinction coefficient for lysozyme at 280 nm was taken as 2.46 cm^2mg^{-1}.

α-synuclein (wild type and mutant A30P) powder was dissolved at concentration 12 mg/ml in phosphate buffer 20 mM at pH 7.4 with 150 mM of added $NaCl$ and 40 μM ThT. The solution was filtered by 0.2 μm *Millipore* syringe filters. Sample concentration was determined by using a calculated extinction coefficient at 280 nm of 0.412 cm^2mg^{-1}. The sample was incubated at 37 $°C$. Fluorescence was measured at 485 nm (excitation at 450 nm)

A.2 PL

Reliability tests

Photoluminescence (PL) measurements were performed on lysozyme samples ($0.18\ mg/ml$) prepared in water at pH 2 (pH was corrected by small amounts of hydrochloride acid). We used a Jasco FP-6500 spectrofluorimeter and a $1\ cm$ length quartz cell. The spectral parameters were set in the following way: excitation wavelength $300\ nm$ for lysozyme intrinsic PL, excitation and emission bandwidth $3\ nm$, scan speed $100\ nm/min$, response $2\ s$. The excitation wavelength for the intrinsic PL experiments was greater than $280\ nm$ to minimize the excitation of tyrosine residues in lysozyme and avoid to damage them. The temperature of the cell holder was regulated by an external circulating bath in the range 60–70 $°C$. Intrinsic photoluminescence of a protein is used to get information about local environment of tryptophans. Emission band of tryptophan can be red-shifted or blue-shifted if the fluorophore is respectively in a more hydrophilic or more hydrophobic environment [133]. So we can know if tertiary structure is changed.

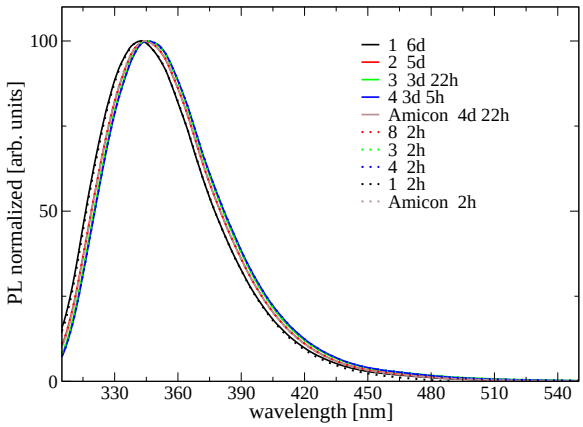

Fig. A.1: Irreversibility of red-shift in the emission spectra (normalized respect to their maximum). Recooling at 25 $°C$ for 2 h (dotted line) or more time (solid line) after incubation at 60 $°C$ for 39 h (black), 67 h (red), 70 h (30 kDa filters) (brown), 95 h (green) and 111 h (blue). Sample concentration: $c = 0.17 mg/ml$

We obtained band peak through a parabolic fit of the maximum. Emission spectra were fit with a set of four Gaussian bands (with centers and widths fixed) to reproduce well the band profile. We calculated the first moment M_1 as a parameter that takes consideration of both centre of band and widthness. It is defined as $\int \lambda P(\lambda) d\lambda$ where $P(\lambda)$ is the normalized profile

with respect to the total emission energy E. In order to test some aspects about emission spectra of PL measurements, we prepared a sample divided into several aliquots. We tried to use also 30 kDa AMICON ultracentrifuge filters to be sure that there were not small pre-existing clusters of proteins. Fig. A.1 shows the emission spectra of lysozyme samples incubated at 60 $°C$ for different amounts of time and thermalized at 25 $°C$ for two hours. They show a red-shift which increases by the incubation time, by manteining the behavior observed at high temperature. Thus the conformational change seems irreversible. We kept some samples at 25 $°C$ for more time, until 6 days and emission spectra are not modified because of waiting. Only a small reduction of the signal can be seen, maybe due to precipitation. So the conformational change is not slowly reversible, but definitively irreversible. This behavior is also reproduced if we use 30 kDa AMICON ultracentrifuge filters to prepare the samples. So small clusters, which potentially could serve as nuclei, have no any off-pathway roles in the lysozyme fibrillation.

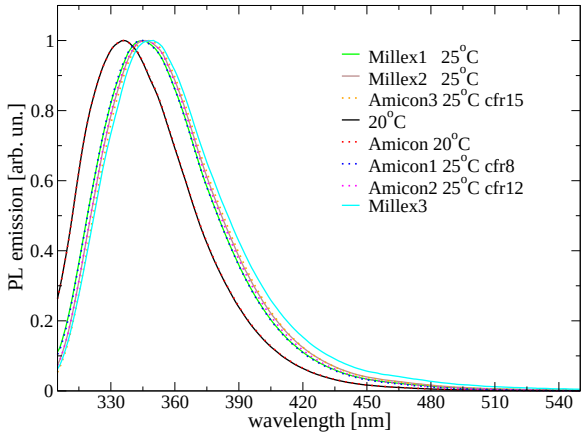

Fig. A.2: Effect of filtration method upon the emission spectra at different incubation times. Samples filtered by *Millex LG* (solid line) or 30 kDa Amicon (dotted line) filters are incubated at 60 $°C$ for 0 h, 67 h, 111 h and 261 h

We investigated better an eventual effect of filtration method upon the emission spectra. We compared samples at 25 $°C$, incubated for the same amount of time at 60 $°C$ by using Millex LG filters (200 nm) and 30 kDa AMICON ultracentrifuge filters. AMICON-filtered samples have raw emission spectra sistematically a little bit higher than the corresponding ones filtered by *Millex LG* filters and the difference increases by incubation time. If we normalize spectra to highlight some differences in the form of the emission profile, we can note that initially there is no difference between two

procedures (Fig. A.2). By increasing the incubation time (this effect is quite evident after 4 days at 60 $°C$) the peak related to AMICON-filtered samples is a bit less red-shifted and sharper. If the effect of two different filters, as we guess, is exclusively to block or not small initial clusters, it seems that monomer is less changed when it is free in solution. This supports that oligomer formation enhances mechanical conformational change of each monomer inside. Anyway, during first three days there is no difference between two filtering methods (at 0.18 mg/ml).

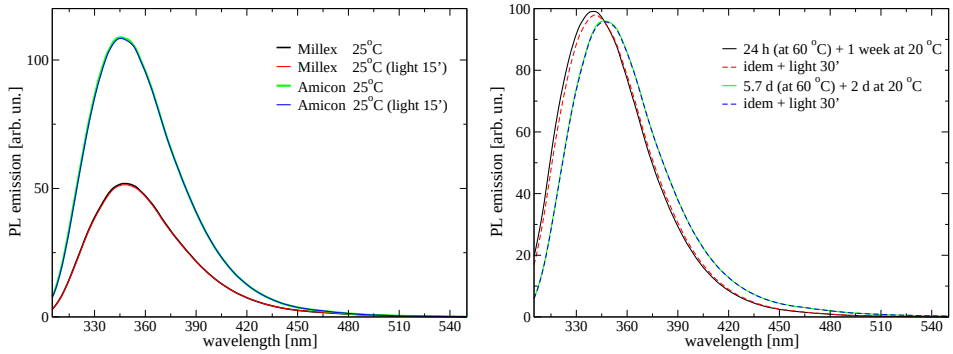

Fig. A.3: a) Effect of lamp energy on samples filtered in two different ways. b) Effect of lamp energy at different stages of the kinetics.

Which is the effect of energy from spectrofluorimeter lamp on emission spectra? Biological fluorophores typically emit because of some movement of the die (see rhodopsin) or of some close chemical groups (see tryptophan) so, if a lot of energy goes on the dye, it could have a mechanical damage. Besides it could be bleached. Fig. A.3 shows that a 15 minutes exposure at 25 $°C$ quenches the signal less than 1 % in both filtering methods, without changes in the form of the band, especially after a long incubation time. At the beginning of the incubation, proteins seem a little bit more sensitive to the lamp energy.

ThT: a fluorescent dye

Thioflavin T (ThT) is a fluorescent dye commonly used for staining *in vivo* and *in vitro* tissues. It was introduced for studying the fibrillogenetic process [256]. ThT is made of a benzothiazole group, a benzene ring and a dimethylamino group (Fig. A.4a) [264, 239, 154]. Spectroscopic properties of ThT are related to the two torsional angles ϕ and ψ.

In aqueous solution and in presence of native proteins, the absorbance spectrum of ThT is characterized by an intense band in the visible region

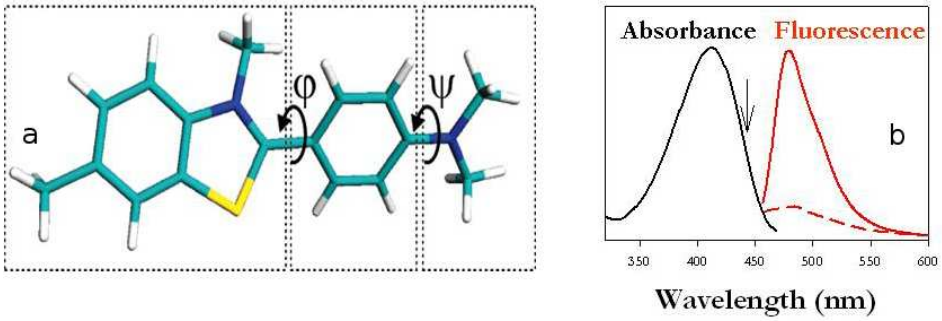

Fig. A.4: a) Structure of ThT with angles ϕ and ψ highlighted. b) Absorption and emission spectra of ThT. Adapted from [76]

at $\approx$ 400–450 nm (Fig. A.4b). Upon excitation at $\approx$ 440–450 nm, ThT has a low emission quantum yield in acqueous solution [165, 139] and a high binding efficiency for fibrils: when bound to amyloid aggregates it shows a bright fluorescence emission in the 475–600 nm region (Fig. A.4b). Moreover, ThT does not affect [172], or only slightly affects, the early stages of fibril formation. Unfortunately there is not a clear quantitative relationship between fluorescence emission and fibril mass, so the specific ThT-fibril binding is ill understood. The reasons for enhanced quantum yield in presence of fibrils are not clear yet, but the most accepted idea is that ThT behaves as a molecular rotor. In its ground state, a dihedral angle ϕ of $\approx 37\,^{\circ}$ between the benzothiazole group and the aminobenzene ring has been reported. Occurrence of the high energy state reflecting a twisted conformation of the ThT molecule ($\phi \approx 90\,^{\circ}$, affects the quantum yield of fluorescence and depends on the physical properties of the enviroment, e.g. viscosity of the solvent [264, 239, 154, 238]. In low viscosity solvents, rotation of the molecule is favoured and no radiative process quences the ThT fluorescence. On the other hand, in viscous solvent and, more generally, in environments where rotation is blocked, the transition is suppressed resulting in an enhanced quantum yield of fluorescence. Such purely theoretical model could explain the bright fluorescence observed in presence of amyloid fibrils. In fact, ThT molecule seems to specifically bind to fibrils, experiencing an environment able to block the transition. However the rotor model does not take into account the specificity of the ThT-fibril binding. Amyloid fibrils structure results to be rich in β-sheet placed perpendicularly to the elongation axis of the fibrils [96, 229]. Moreover, a number of studies suggests a geometrical arrangements of the β-sheet chains leading to a sort of empty cavity generally

referred in literature as β-sheet channel [96, 95, 229, 121]. Krebs et al. [128] proposed that enhanced fluorescence can only occur when ThT molecule is bound in a highly directional mode (Fig. A.5).

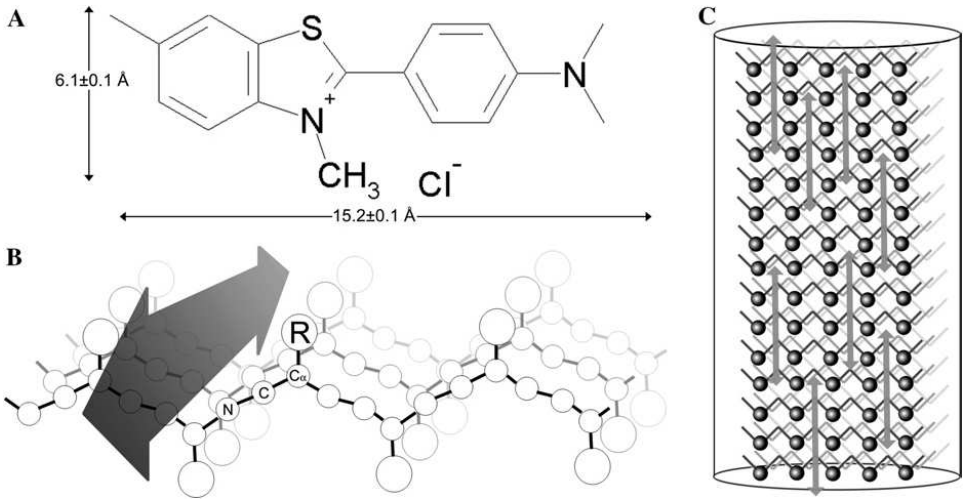

Fig. A.5: a) Structure of ThT. b)-c) Two possible binding sites for ThT on β-sheet. Adapted from [128]

ThT could bind to such cavity with its long axis parallel to the elongation axis of the fibril. Furthermore, different morphologies of amyloid fibrils can not only lead to different ThT binding kinetics, but they may also induce different ThT fluorescence quantum yields [180]. As pointed out by Manno and co-workers, in general capability and specificity of ThT-fibrils binding need to be verified for each specific system [151]. In particular, the available accessible fibrils surface is a crucial point to evaluate the goodness of the ThT assay. For example, in the case of thin and dispersed fibrils, ThT signal can be easily related to the fibril concentration. In contrast, when fibrils start to assemble forming bundles, ratio between mass and available surface of fibrils dramatically increases, leading to an erroneous evaluation of fibril concentration [151]. Other possible drawbacks for a proper ThT staining can be due to the ThT-fibril electrostatic interaction. Pedersen and coworkers showed that ThT quantum yield can also change dependent on different concentration of GdmCl in the solvent, being such effect possibly due to the charge screening by chloride ions between the positive charged ThT ions and fibril [180]. We used ThT only in neutral conditions, where ThT is discharged, because at acidic pH its binding efficiency is very low. This is probably due to the high electrostatic repulsion between ThT and lysozyme.

A.3 DSC

Differential Scanning Calorimetry (DSC) data were obtained by using two different *power compensation* calorimeters for suitable range of scan rate. An *Hart Scientific* calorimeter (mod. 3705/06) coupled with a *transputer* driven control unit [29] was used for measurements at slow scan rates (0.07–0.3 $°C\ min^{-1}$). A *Perkin-Elmer* calorimeter (*Pyris Diamond*), coupled with a *Cryofill* device, was used for measurements at higher scan rates (5–40 $°C/min$). The first one was calibrated by naphthalene to correct the peak temperature [111]; the second one was calibrated by indium to correct peak temperature and heat flow. 270 μl of sample were introduced in a 1 ml steel cell for *Hart Scientific* apparatus and 20 μl of sample in a 20 μl steel cell for *Perkin-Elmer* apparatus. The same volume of buffer solution was inserted in a similar reference-cell. In both equipments the cells were kept under nitrogen laminar flow to avoid air-water condensation. Measures were performed on concentrated samples (10–60 mg/ml) in a temperature range of 25–85 $°C$. The specific heat Δc_p is calculated from the raw calorimetric signal $S(T)$ according Eq. A.1:

$$c_p = -\frac{\partial H}{\partial T}\bigg|_p = -\frac{K(T)S(T)}{v} \tag{A.1}$$

where $K(T)$ is the calibration curve and v is the scan rate. Sometimes the plots show c_p that is the heat capacity (for mass unit) of protein solution with respect to buffer solution. It represents the net contribution of the protein. We calculated c_p^{exc} by subtracting a quadratic baseline, which is obtained by fitting the calorimetric curve without the peak. Typically a DSC thermogram is like in Fig. A.6: where T_m is the unfolding temperature, the area under the curve $c_p^{exc}(T)$ is the experimental enthalpy change ΔH_{exp}, and Δc_p indicates the change of freedom degrees from native to unfolded state.

We can calculate the functions $\Delta H(T)$ and $\Delta S(T)$ from $c_p(T)$ or the most important calorimetric parameters (ΔH^{exp}, T_m, Δc_p) through the Kirchoff equations (Eq. A.2-A.3) [196]:

$$\Delta H(T) = \int_{T_0}^{T} c_p\, dT = \Delta H^{exp} - \Delta c_p(T_m - T) \tag{A.2}$$

$$\Delta S(T) = \int_{T_0}^{T} \frac{c_p}{T}\, dT = \frac{\Delta H^e xp}{T_m} - \Delta c_p ln\left(\frac{T_m}{T}\right) \tag{A.3}$$

The Gibbs-Helmontz equation allows to calculate ΔG between native and

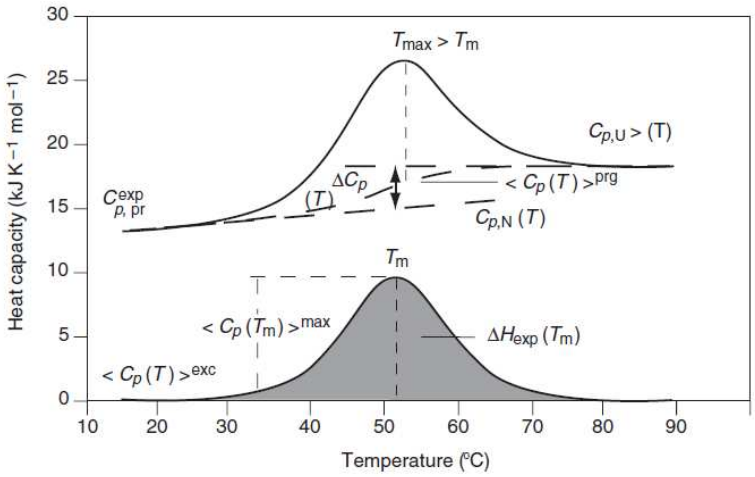

Fig. A.6: Typical DSC thermogram, effect of buffer subtraction and the most important thermodynamic informations collectable from easy analysis of experimental data.

unfolded states [134]:

$$\Delta G(T) = \Delta H(T) - T\Delta S(T) \tag{A.4}$$

A.4 SAXS

Measures about mutant α-Sn were collected at X33 sinchrothron's beamline in Hamburg. Small Angle X-ray Scattering (SAXS) is a very powerful technique to get informations about structure of solution and form factor of objects in it. It is a non invasive method to study proteins in solution. Fig. A.7 shows the working principle. An X-ray beam, suitably collimated, is directed on a protein sample and scatters because of interaction with electronic cloud of atoms in the protein. The coherence properties of incident radiation generate a diffraction pattern which gives informations about relative distances among atoms in the protein.

If N identical objects are present in solution, the general expression for scattering intensity $I(\mathbf{q})$is:

$$I(\mathbf{q}) = \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} f_i(\mathbf{q}) f_j(\mathbf{q}) exp\left[i\mathbf{q}(\mathbf{r}_i - \mathbf{r}_j)\right] \right\rangle \tag{A.5}$$

where $f_i(\mathbf{q})$, $\mathbf{q}$, $\mathbf{r}_i$ are repectively the scattering factor of the objects, the

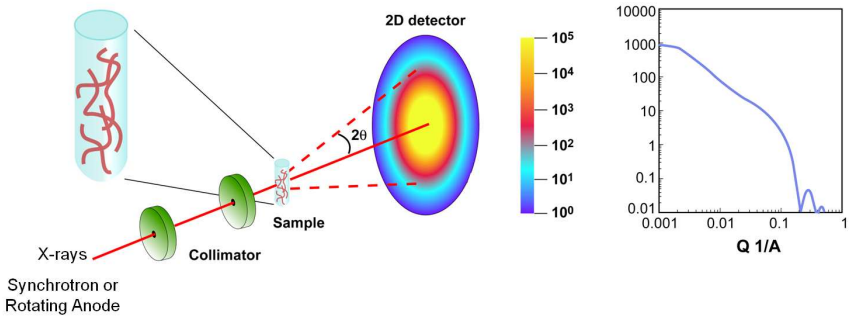

Fig. A.7: Working principle of SAXS. An incident X-ray beam interacts with proteins and produces a diffraction pattern on a detector. 2θ is the scattering angle.

scattering vector (Eq. A.6) and the position of i^{th} object with respect to a fixed reference system. The symbol <...> indicates the ensemble average.

$$q = \frac{4\pi sin\theta}{\lambda} \tag{A.6}$$

If we assume that the objects are spherical, we can use the *decoupling approximation*:

$$I(q) = NP(q)S(q) \tag{A.7}$$

where $P(q)$ is the form factor of the object and $S(q)$ is the structure factor of the solution. This approximation can be used also for not spherical objects and a not monodispersed solution. When concentration is quite low, as in our case, we can approximate $S(q) = 1$ and the scattered intensity is simply:

$$I(q) = \int_0^\infty m^2(R)D(R)P(R,q)\,dR \tag{A.8}$$

where $m(R)$, $P(R,q)$ and $D(R)$ are respectively the number density of particles, a function depending by dimensionality of the system, form factor of objects R sized and the particle number density with size between R and $R + d(R)$. Therefore the key concept of these calculations is the additivity of scattering intensity. If we have different species in solution, each one is characterized by a scattering pattern and the total collected scattered intensity is the sum of all contributions, suitably weighted for their concentration. Determination of the form factor of the species and their relative amounts is not a simple operation. If we know all form factors except one and fix some constraints on relative concentrations of species and on accuracy of

fit, a minimization criterium is sufficient to solve our problem. Therefore we must know how many components are present in solution and the form factor of all species, except that one we want to determine. Of course the method is accurate if polidispersity of each species is small. The analysis is quite complex and some parameters, extracted directly from raw data can be a good guideline. In particular, the behavior of the scattering intensity at very small values of momentum transfer (q $\rightarrow$ 0) is strictly related to overall particles parameters [241]:

$$I(q) = I(0)\left[1 - \frac{1}{3}R_G^2 q^2 + O(q^4)\right] \approx I(0)exp\left(-\frac{1}{3}R_G^2 q^2\right) \tag{A.9}$$

where R_G is called *Guinier radius* and $I(0)$, scattered intensity at zero angle, is strongly related to the volume V of particles:

$$I(0) = 4\pi \int_0^{D_{max}} P(r)\, dr = (\Delta\rho)^2 V^2 \tag{A.10}$$

where $P(r)$ is the Fourier transform of $I(q)$, $\Delta\rho$ is the excess scattering length density and D_{max} is the maximum intraparticle length. Thus $I(q)$ has a linear dependence by q^2 in a log-log plot: *Guinier region*.

A.5 SALS

While X-ray scattering signal is produced by the difference in electronic density, light scattering signal is due to the difference in the refraction index. Also the difference in the range of wavelength leads to observe different things by these two techniques: SAXS allows to have detailed informations on small scale, at molecular level, so we can study intermolecular mean force potential (at high concentration) or geometric properties of single molecules (at low concentration). On the other hand, light scattering gives less informations about intermolecular interactions and geometric properties of the molecules, but it allows to study more deeply the aggregation processes on micro- and meso-scale.

Small-Angle Light Scattering (SALS) measurements were performed by an home-made instrument. Fig. A.8 shows its block diagram.

The source is a 30 mW *Melles-Griot* He-Ne laser (λ = 632.8 nm) and we used a *Pulnix* TM765 CCD camera. The system was built following published work [70]. In particular the laser beam, vertically polarized is aligned to a glass S and attenuated by a neutral filter A at variable density.

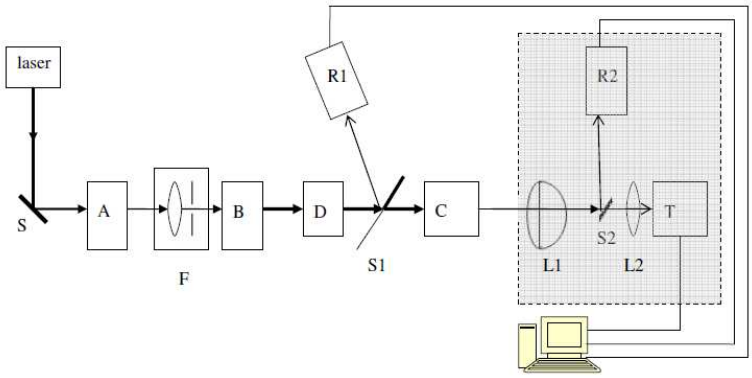

Fig. A.8: Block diagram of SALS equipment.

Laser beam crosses a spatial filter F (convergent lens and a 100 μm size pinhole). The beam is enlarged through the beam expander B up to 3 mm. Diaphragm D is used to select first *Fresnel* zone of the beam. Part of that beam is reflected by glass S_1 (towards the detector R_1); the rest crosses the sample. The transmitted beam enters in a dark room, where the acromatic twin lenses L_1 provides for Fourier transformation. A small glass S_2 excludes the stray light from the beam and avoids to saturate the detector camera T. Measurable intensities range from 1.0 to 3 10^5 in arbitrary units, thanks to a software integration of multiple exposure times 1/60–1/10000 s. After background subtraction, reliable measurements could be made over a range of scattering angles of 0.1°–11°(corresponding to scattering vector magnitudes of 0.02–2 μm^{-1}) with a maximum resolution of 0.01° [30]. The measurements were done into quartz cuvettes with an optical path of 2 mm, at room temperature. In order to avoid *multiple scattering* in aggregated samples, we diluted the sample with the buffer solution to obtain a transmittance $\geq 90\%$. We checked that the profile of the structure functions was not altered by successive dilutions, once multiple scattering was supressed. Very high dilutions were avoided, since a loss of material was detected (data not shown).

Preliminary measurements at room temperature were performed on incubated samples at 81 $°C$ (until it reaches the final plateau, see Fig. 2.16) before measuring structure functions. In fact the samples are turbid, so we need to diluite them to avoid multiple scattering. However we must control that diluition does not break the aggregates and does not modify strongly their morphological properties. Fig. A.9 shows different diluitions (1:2, 1:4, 1:8, 1:16) of samples by using the same buffer.

We observe two features at all concentrations. In concentrated samples

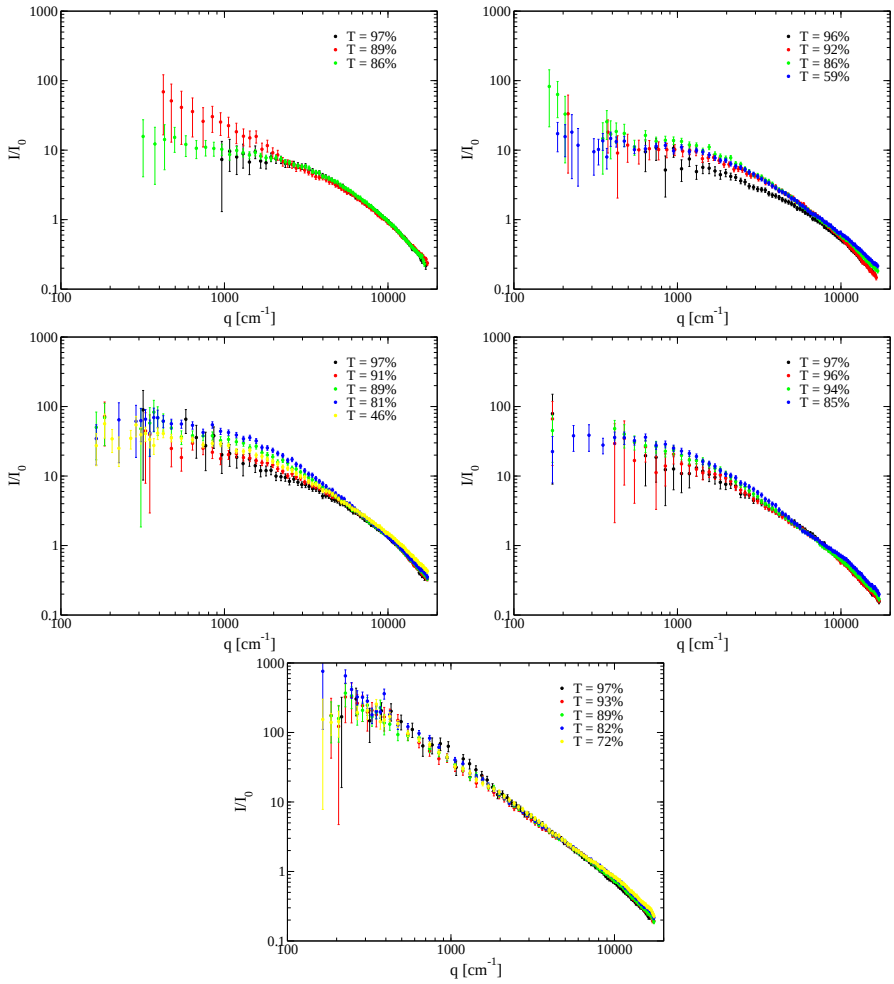

Fig. A.9: Different diluitions of sample incubated at 81 $°C$: a) 0.23 mg/ml; b) 0.50 mg/ml; c) 0.71 mg/ml; d) 0.94 mg/ml; e) 1.92 mg/ml.

the structure function is higher at high q values, while in diluted ones which have trasmittivity more than 95 % the structure function bends at low q values. The first effect is due to multiple scattering; the second one is due to the lost of some material from the aggregates. This suggests that the macroscopic aggregates are not bound by strong interactions, as viscosity measurements confirm (Fig. 2.22).

A.6 DLS

Dynamic Light Scattering (DLS) measurements were performed by a Brookhaven Instrument. Samples were placed in a thermostated cell compartment of a

Brookhaven Instruments BI200-SM goniometer equipped either with 100 mW Ar laser tuned at 514.5 nm or with a He-Ne laser tuned at 632.8 nm. The temperature was controlled within 0.05 $^{\circ}C$ by using a thermostated recirculating bath. The scattered light intensity, at 90°, and its time autocorrelation function $g_2(t)$ were measured by using a *Brookhaven* BI-9000 correlator. A photodiode positioned at 0° was used to collect forward scattered light and to measure the sample turbidity. The scattered intensity is normalized with respect to toluene, whose Rayleigh ratio at 632.8 nm and 514.5 nm was taken respectively as $14.0 \cdot 10^6 cm^{-1}$ and $32.0 \cdot 10^6 cm^{-1}$. Autocorrelation functions $g_2(t)$ can be fitted by a discrete expression (Eq. A.11):

$$g_2(t) = 1 + \beta \left[\sum_i A_i exp(-t/\tau_i) + A_0 \right]^2 \tag{A.11}$$

where the summation is extended to all species i (monodispersed for hypothesis) guessed to be in solution, τ_i is the correlation or relaxation time of i^{th} diffusional process. It can be shown that for a diffusional process, $\tau_i^{-1} = D_i q^2$, where D_i is the apparent diffusion coefficient of i^{th} species and q is the scattering vector (which depends upon the scattering angle θ by the relation $q = 4\pi n_0 \lambda^{-1} sin(\theta/2)$, where n_0 is the medium refractive index [24]. A_i is the relative weight of each species and A_0 is a constant background which considers for objects with very slow diffusional process. β is an instrumental factor which considers the pin hole. A more detailed model includes a polidispersity of the species by using the cumulants method [127].

$$g_2(t) = 1 + \beta \left[\sum_i A_i exp\left(-t/\tau_i + \frac{\mu_2^i t^2}{2!} - \frac{\mu_3^i t^3}{3!} + ... \right) \right]^2 \tag{A.12}$$

where μ_2^i and μ_3^i are respectively the second and the third momentum of distribution $G(\Gamma)$ with $\Gamma = \tau_i^{-1}$.

However another method to take into account the polidispersity of the species is more stable in the fitting procedure [82]:

$$g_2(t) = B + \beta \left[\sum_i A_i exp(-t/\tau_i)(1 + \frac{\mu_2^i}{2!} t^2 - \frac{\mu_3^i}{3!} t^3) \right]^2 \tag{A.13}$$

where B is the baseline. In the most complex systems, characterized by a quite high polidispersity the best way to analyze the correlation functions is to use $CONTIN$. It is an algorithm of regularization which calculates the inverse Laplace transform of $g_1(t)$ by using two criteria: the minimization of

chi square in fitting procedure of $g_1(t)$ and the regularization of the distribution through the research of a solution as smoothed as possible. The result is the distribution of diffusion times [198, 199], related to size distribution through Stokes-Einstein relation (Eq. 4.9).

$$g_2(t) = 1 + \beta \left[\int_0^\infty P(D_h) exp(-D_h q^2 t) dD_h \right]^2 \tag{A.14}$$

It is the continuous extension of Eq. A.11.

A.7 Rheology

Rheology studies the visco-elastic properties of a material. An ideal solid is elastic and we define the rigidity module G_0 as:

$$G_0 = \frac{\tau}{\gamma} \tag{A.15}$$

where γ (*shear strain*) is the mechanical response of τ (*shear stress*). An ideal fluid is viscous and we define the viscosity η_0 as:

$$\eta_0 = \frac{\tau}{\dot{\gamma}} \tag{A.16}$$

where $\dot{\gamma}$ is called *shear rate*. Real matter has both behaviors. In particular, if we apply a stress such that shear strain is on the same scale length of intermolecular bonds, the sample will show an elastic behavior. On the contrary, if we apply a stress such that shear strain is on a larger scale length, the bonds will be broken and the sample will show a viscous behavior [18]. We used the Maxwell model to sum the both effect: it consists in the sum of elastic and viscous strains (Eq. A.17).

$$\dot{\gamma}(t) = \frac{\dot{\tau}(t)}{G_0} + \frac{\tau(t)}{\eta_0} \tag{A.17}$$

In the frequence domain Eq. A.17 is equivalent to Eq. A.18, where G' and G'' are related respectively to elastic and viscous component; ω is the angular oscillation frequency.

$$G(\omega) = \frac{\tau(\omega)}{\gamma(\omega)} = \left(\frac{1}{G_0} - \frac{i}{\eta_0 \omega} \right)^{-1} = G' + iG'' \tag{A.18}$$

We used a *stress-controlled* rheometer (AR1000, *T.A. Instruments UK*) to perform the measurements. Samples incubated at 81 $^\circ C$ were pipetted onto

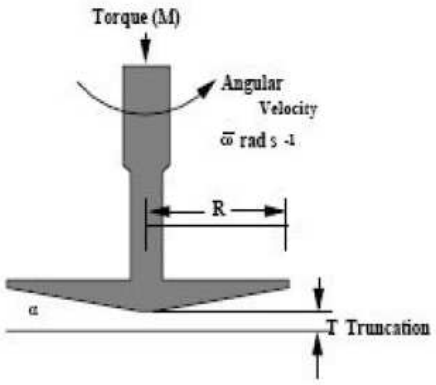

Fig. A.10: Geometry (cone) used for rheological measurements.

the plate, thermally controlled by a Peltier device. The geometry used was a 40 mm titanium cone with an angle of less than 1° and a 26 μm truncation length (*T.A Instruments*) (Fig. A.10).

The advantage of this geometry is to apply a shear stress independent from radial distance. The detection of angular displacement is made by an optical device (sensitivity 0.6 μrad) Any excess sample was blotted away with filter paper, and paraffin oil (GPR grade, BDH, UK) was layered around exposed regions. A solvent-trap cover-slip was also used to prevent evaporation. Viscoelastic spectra were measured at 20 $°C$ in the range 0.1–200 $rad\ s^{-1}$, with a strain of 0.01. Viscosities were also measured at 20 $°C$ as a function of shear rate (from 1 to 1000 s^{-1}).

A.8 CD

Circular dichroism measures the different absorption of left and right (A_L and A_R) circularly polarized light (Eq. A.19). This phenomenon is due to different coefficient for the two components and characterizes chiral molecules. The chromophore is the diedre angle in the peptidic bond (note "$N_{res} - 1$" in Eq. A.20).

$$\Delta\epsilon = \frac{A_L - A_R}{cL} \tag{A.19}$$

where $\Delta\epsilon$, c and L are respectively the difference of dichroic extinction coefficient, sample concentration and the pathlength [68]. We used far UV circular dichroism to get information about secondary structure. Every secondary structure has a characteristic CD spectrum, so in principle we could

know their relative amounts. Typical secondary structures (α-helix, β-sheet and *random coil*) are represented in Fig. A.11 [93]. They were obtained by poly-L-lysine chains at different pH conditions. In particular poly-L-lysine is *random coil* at pH 7, α-helix at pH 10.8 and β-sheet at pH 11.1 (after heating to 52 $°C$ and recooling).

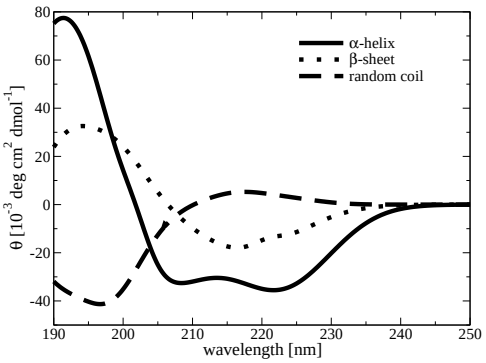

Fig. A.11: Mean residual ellipticity from far-UV CD spectra of typical secondary structures. These spectra are obtained putting poly-L-lysine in different pH and temperature conditions to have 100 % α-helix (solid line), 100 % β-sheet (dotted line) or 100 % *random coil* (dashed line).

We used two instruments; first one is a *Jasco* J-815 spectrometer and set the following spectral parameters: bandwidth 1 nm, scan speed 10 nm/min, response time 4 s. We also used an *OLIS* (DSM 10 CD) spectrometer and set slitwidth 0.5 mm (corresponding to bandwidth 2 nm). The acquisition is made in *step-mode* (at fixed wavelengths) and the response time is automatically adjusted during the measurements to have a good signal to noise ratio. The temperature was set by a computer-controlled Peltier device in both equipments. Each spectrum was obtained by subtracting the corresponding solvent spectrum and reported as Mean Residual Ellipticity (θ).

$$\theta_{deg\ cm^2\ dmol^{-1}} = \frac{\theta_{mdeg} P.M.}{10(N_{res}-1)d_{cm}c_{mg/ml}} \tag{A.20}$$

where θ_{mdeg}, $P.M.$, N_{res}, d and c are respectively the CD raw signal, the molecular mass and the number of amino acids of the protein, the pathlength of the cell and the concentration of the sample.

Therefore we chose to collect the measurements at protein concentration 0.21 mg/ml and use a quartz cell with 0.1 cm pathlength. In order to analyze CD spectra (in the range 178–260 nm) we used the software package by *OLIS*. It allows to choose different algorithms to deconvolve the signal into

secondary structure components. In particular we used $CONTINLL$, which fits the CD signal of unknown proteins by a linear combination of the spectra of a large basis of reference proteins with known conformations. [200]. In this method the contribution of each reference spectrum is kept small unless it contributes to a good agreement between the theoretical best fit curve and the raw data. A good accuracy is obtained in the estimate of β-turn content. We used a reference set made of 48 proteins (42 soluble and 6 unfolded). The analysis returned the content of α-helix, α_{3-10}-helix, β-turn, *proline*-turn and *random coil*.

A.9 AFM

At the end of the turbidity experiments the incubated solution was diluted 100 times (10 μl of sample into 1 ml buffer solution, quenched to room temperature), or 10000 times and used for atomic force microscopy (AFM) experiments. A few microliters of the solution were dropped onto a freshly cleaved mica substrate. After a few minutes, the sample was washed dropwise with *Millipore Super-Q* water, and then dried with a gentle stream of dry nitrogen. The drying is expected to remove most of the surface water, while keeping the protein aggregates locally hydrated. Images of the protein aggregates were recorded with a *Multimode Nanoscope V* Atomic Force Microscope (*Veeco Instruments, Santa Barbara, CA, USA*), operating in *tapping mode*. We used rigid cantilevers (model RFESP) with resonance frequencies of about 85 kHz, and equipped with single crystal silicon tips with a nominal radius of curvature of 5 to 10 nm. The used resolution of images is (512×512 points), and line scan rate was 0.5 Hz. The analysis of AFM images was done by using *Gwyddion*, a free and Open Source software (http://gwyddion.net/) developed by support of Czech Metrology Institute. We perform the measurements in tapping mode. The experimental principle is based on the forced and damped oscillator. The cantilever oscillates and a tip is located on the end of it (Fig. A.12b). Reflections of a laser beam on cantilever are revealed by a detector and the instrument calculates the height of the objects deposited on the mica' surface by analizing the dephasing that object-cantilever interaction produces on the cantilever (Fig. A.12a). We used this material because it can be reduced in sheets and it can charge negatively. This technical aspect is very convenient because lysozyme is positively charged in our solvent conditions, so protein remains electrostatically anchored to the mica' surface and cannot be shifted by cantilever movements during the measurements. Typically we must dilute the sample before performing the AFM measurements because the number of

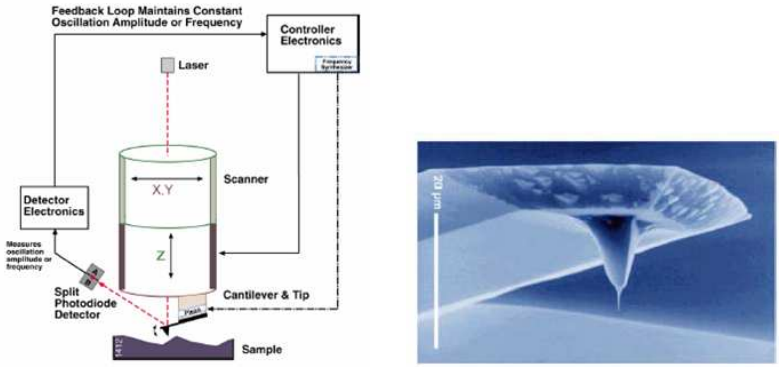

Fig. A.12: a) Principle of function of AFM. b) Zoom of a typical tip for AFM imaging.

objects should not be too high on the mica's plate. In particular we diluted our incubated sample (about 18.5 mg/ml) 100 or 10000 times. The best dilution is 10000 times (2 $\mu g/ml$) because, by diluting only 100 times, we observe still a network, caused by deposition of a great number of fibrils. We put a drop (about 20 μl) of the sample on a mica substrate and dry the sample by gentle nitrogen flux. When we analyze the AFM images we must take into account that the tip mounted on the cantilever has a finite size. Thus the sides of the tip start to interact with the object before that the tip is next to it. This effect is as big as bigger is the curvature radius of the tip and we "see" the object larger than it is. If we used different tips, an object seems to have different widths, but the height is quite the same. In tapping mode the height is also affected of an error because the tip oscillates as closer as possible to it, without touching it (if the sample has not too irregularities!). The apparent height is influenced from tip-object interaction and we must take into account that the sample is not completely dry, but it is covered by an hydration shell. Thus raw AFM images can give only qualitative informations. We need to provide deconvolution of the images to have some quantitative parameters about them. We used a standard formula (Eq. A.21) for parabolic tip on an isolated object:

$$w' = w - 2(2hR)^{0.5} \tag{A.21}$$

where w, w', h and R are respectively apparent and deconvolute width, apparent height of the object and curvature radius of the tip. The tips that we used, VEECO RFESP and Nanosensor PPP-NCHR-50, have respectively curvature radius of 8 and 7 nm [186].

Appendix B

PhD activities

Schools

- International School of biophysics. 36th Course: Multidimensional optical fluorescence microscopy towards nanoscopy. Erice, April 19–29 (2008)
- International Summer School of biophysics. 10th Course: Supramolecular structure and function. Croatia (Red Island) Sept. 19 – Oct 01 (2009)

Exams

- Laboratorio di ottica (Prof. M. Cannas)
- Interazione radiazione-materia (Prof. R. Passante)
- Small Angle X-ray Scattering: teoria e metodi di analisi (Prof. M. Leone)
- Fisica statistica (Prof. M. Palma)

Posters and oral communications

Raccosta S., Bulone D., Giacomazza D., Leone M., Martorana V. and Manno M. Thermal irreversible unfolding and aggregation of Lysozyme. - International School of biophysics. 36th Course: Multidimensional optical fluorescence microscopy towards nanoscopy. Erice, April 19–29 2008 (oral)

Raccosta S., Manno M., Bulone D., Giacomazza D., Leone M., Martorana V. and San Biagio P.L. Thermal irreversible unfolding and aggregation of Lysozyme. SIBPA. Roma Settembre 2008 (oral)

Raccosta S., Martorana V. and Manno M. Fibrillogenesis of Hen Egg-White Lysozyme at acidic pH. International Bunsen Discussion Meeting, Halle an der Saale (Germany) February 08–11 (2009)

Raccosta S., Martorana V., Manno M. and San Biagio P.L. Fibrillogenesis of hen egg-white lysozyme at acidic pH. EBSA, Genova, July 11–15 (2009)

Raccosta S., Martorana V., Manno M. and San Biagio P.L. Fibrillogenesis of hen egg-white lysozyme at acidic pH. 13th ECSBM, Palermo, Aug 28 – Sept 02 (2009)

Raccosta S., Martorana V., Manno M. and San Biagio P.L. Fibrillogenesis of hen egg-white lysozyme at acidic pH. International Summer School of Biophysics, Rovinj (Red Island) - Croazia, Sept 19 âĂŞ- Oct 01 (2009)

Raccosta S., Martorana V. and Manno M. Role of repulsive interaction on the fibrillogenesis of hen egg-white lysozyme. 20th SIBPA, Arcidosso, Sept 11–14 (2010)

D'Amico M., Cannas M., Martorana V., Raccosta S. and Manno M. Preferential solvatation of TFE in Lysozyme: a luminescence study. 13th ECSBM, Palermo, Aug 28 – Sept 02 (2009)

Publications

Raccosta S., Manno M., Bulone D., Giacomazza D., Militello V., Martorana V. and San Biagio P.L. Irreversibile gelation of thermally unfolded proteins: structural and mechanical properties of lysozyme aggregates. European Biophysics Journal. 39 (6), 1007–1017, 2010

D'Amico M., Raccosta S., Cannas M., Martorana V. and Manno M. The existence of metastable intermediate of lysozyme conformation highlights the role of alcohols in altering protein stability. Submitted to J. Phys. Chem. B

Raccosta S. et al. Aggregation processes of Hen Egg-White Lysozyme induced by thermal denaturation (To be submitted to Biophys. Chem.)

Raccosta S., Martorana V. and Manno M. Lysozyme fibrillogenesis: competitive interactions and oligomer formation. (In preparation)

Stages

Stage at Department of Medicinal Chemistry (Faculty of Pharmaceutical Sciences) in Copenhagen for advanced analysis on SAXS data. Study on mutant α-Synuclein A30P. Duration: 6 months (March–August 2010).

Other

Assistance to training activity upon Differential Scanning Calorimetry.

Development of a Fortran code for the simulation of X-ray structure functions with known PDB file.

Scientific animator at "Festival della Scienza" in Palermo, (Loggiato San Bartolomeo) Dec 04 – Jan 02 2009-10.

Scientific animator for project "Esperienza insegna 2010" in Palermo, (Polo Didattico) Feb 18–25 2010

Lab-demonstrator for the project "Diffusione della cultura scientifica" at Institute of Biophysics (CNR) in Palermo, May–Nov 2010

Bibliography

[1] K. Ahrer, A. Buchacher, G. Iberer, and A. Jungbauer. Thermodynamic stability and formation of aggregates of human immunoglobulin g characterised by differential scanning calorimetry and dynamic light scattering. *J. Biochem. Biophys. Methods*, 66:73–86, 2006.

[2] S. Alexander, P.M. Chaikin, P. Grant, G.J. Morales, P. Pincus, and D. Hone. Charge renormalization, osmotic pressure, and bulk modulus of colloidal crystals: Theory. *J. Chem. Phys.*, 80(11):5776–5781, 1984.

[3] M. Amani, A.A. Moosavi-Movahedi, G. Floris, A. Mura, B.I. Kurganov, F. Ahmad, and A.A. Saboury. Two-state irreversible thermal denaturation of euphorbia characias latex amine oxidase. *Biophys. Chem.*, 125(2-3):254–259, 2007.

[4] C. B. Andersen, M. Manno, C. Rischel, M. Thórólfsson, and V. Martorana. Aggregation of a multidomain protein: A coagulation mechanism governs aggregation of a model igg1 antibody under weak thermal stress. *Protein Sci.*, 19:279–290, 2010.

[5] J.M. Andrews and C.J. Roberts. A lumry-eyring nucleated polymerization model of protein aggregation kinetics: 1. aggregation with pre-equilibrated unfolding. *J. Phys. Chem. B*, 111(27):7897–7913, 2007.

[6] S. Arai and M. Hirai. Reversibility and hierarchy of thermal transition of hen egg-white lysozyme studied by small-angle x-ray scattering. *Biophys. J.*, 76:2192–2197, 1999.

[7] L.N. Arnaudov and R. de Vries. Thermally induced fibrillar aggregation of hen egg white lysozyme. *Biophys. J.*, 88(1):515–526, 2005.

[8] D. Arosio, F. Ricci, L. Marchetti, R. Gualdani, L. Albertazzi, and F. Beltram. Simultaneous intracellular chloride and ph measurements using a gfp-based sensor. *Nature Methods*, 7(7):516–U44, 2010.

[9] A. Arroyo-Reyna, S.R. Tello-Solis, and A. Rojo-Domínguez. Stability parameters for one-step mechanism of irreversible protein denaturation: a method based on nonlinear regression of calorimetric peaks with nonzero δc_p. *Anal. Biochem.*, 328(2):123–130, 2004.

[10] M. Aubouy, E. Trizac, and L. Bocquet. Effective charge versus bare charge: an analytical estimate for colloids in the infinite dilution limit. *J. Phys. A - Math. Gen.*, 36(22):5835–5840, 2003. International Conference on Strongly Coupled Coulomb Systems (SCCS), SANTA FE, NEW MEXICO, SEP 02-06, 2002.

[11] A.I. Azuaga, C.M. Dobson, P.L. Mateo, and F. Conejero-Lara. Unfolding and aggregation during the thermal denaturation of streptokinase. *Eur. J. Biochem.*, 269:4121–4133, 2002.

[12] J.K. Baird and Y. Kim. Theory of the nucleation of protein macromolecular ions. *Mol. Phys.*, 100(11):1855–1866, JUN 2002.

[13] H. Bajaj, V.K. Sharma, and D.S. Kalonia. Determination of second virial coefficient of proteins using a dual-detector cell for simultaneous measurement of scattered light intensity and concentration in sec-hplc. *Biophys. J.*, 87(6):4048–4055, 2004.

[14] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, and J.A. McCammon. Electrostatics of nanosystems: Application to microtubules and the ribosome. *PNAS USA*, 98(18):10037–10041, 2001.

[15] R.L. Baldwin. How hofmeister ion interactions affect protein stability. *Biophys. J.*, 71(4):2056–2063, 1996.

[16] T. Banerjee and N. Kishore. A differential scanning calorimetric study on the irreversible thermal unfolding of concanavalin a. *Thermochimica Acta*, 411:195–201, 2004.

[17] L. Bao, S. Chatterjee, S. Lohmer, and D. Schomburg. An irreversible and kinetically controlled process: Thermal induced denaturation of l-2-hydroxyisocaproate dehydrogenase from lactobacillus confusus. *The Protein Journal*, 26:143–151, 2007.

[18] H. A. Barnes, J. F. Hutton, and K. Walters. *An Introduction to Rheology.* Elsevier Science B. V., Amsterdam, 1989.

[19] G. Barone, C. Giancola, and A. Verdoliva. Dsc studies on the denaturation and aggregation of serum albumins. *Thermochimica Acta*, 199:197–205, 1992.

[20] D.C. Bas, D.M. Rogers, and J.H. Jensen. Very fast prediction and rationalization of pka values for protein-ligand complexes. *Proteins*, 73:765–783, 2008.

[21] K. Baussay, C.L. Bon, T. Nicolai, D. Durand, and J.P. Busnel. Influence of the ionic strength on the heat-induced aggregation of the globular protein *b*-lactoglobulin at ph 7. *Int. J. Biol. Macromol.*, 34:21–28, 2004.

[22] B.M. Baynes and B.L. Trout. Rational design of solution additives for the prevention of protein aggregation. *Biophys. J.*, 87(3):1631–1639, 2004.

[23] W. J. Becktel and J. A. Schellman. Protein stability curves. *Biopolymers*, 26:1859–1877, 1987.

[24] B. J. Berne and R. Pecora. *Dynamic Light Scattering.* Wiley.Interscience, New York, NY, 1976.

[25] F. Bonneté, S. Finet, and A. Tardieu. Second virial coefficient: variations with lysozyme crystallization conditions. *J. Cryst. Growth*, 196(2-4):403–414, 1999. 7th International Conference on Crystallization of Biological Macromolecules, GRANADA, SPAIN, MAY 03-08, 1998.

[26] Michael L. Broide, Tina M Tominc, and Mark D. Saxowsky. *Phys. Rev. E*, 53:6325–6335, 1996.

[27] M. Brucale, M. Sandal, S. Di Maio, A. Rampioni, I. Tessari, L. Tosatto, M. Bisaglia, L. Bubacco, and B. Samori. Pathogenic mutations shift the equilibria of alpha-synuclein single molecules towards structured conformers. *Chem. Bio. Chem.*, 10(1):176–183, JAN 5 2009.

[28] M. Buck, H. Schwalbe, and C. M. Dobson. Characterization of conformational preferences in a partly folded protein by heteronuclear nmr spectroscopy: Assignment and secondary structure analysis of hen egg-white lysozyme in trifluoroethanol. *Biochemistry*, 34:13219–13232, 1995.

[29] D. Bulone, L. G. Fornili, S. L. Fornili, M. Lapis, and P. L. San Biagio. Transputerbased upgrading of a differential scannino calorimeter. *Meas. Sci. Techmol.*, 5:1443–1447, 1994.

[30] D. Bulone, D. Giacomazza, V. Martorana, J. Newman, and P. L. San Biagio. Ordering of agarose near the macroscopic gelation point. *Phys. Rev. E*, 69:0411401:1–9, 2004.

[31] V. E. Bychkova, A. E. Dujsekina, S. I. Klenin, E. I. Tiktopulo, V. N. Uversky, and O. B. Ptitsyn. Molten globule-like state of cytochrome c under conditions simulating those near the membrane surface. *Biochemistry*, 35:6058–6063, 1996.

[32] G. Caldarelli and P. De Los Rios. Cold and warm denaturation of proteins. *Biol. Phys.*, 27:229–241, 2001.

[33] X. Cao, J. Li, X. Yang, Y. Duan, Y. Liu, and C. Wang. Nonisothermal kinetic analysis of the effect of protein concentration on bsa aggregation at high concentration by dsc. *Thermochimica Acta*, 467:99–106, 2008.

[34] F. Cardinaux, A. Stradner, P. Schurtenberger, F. Sciortino, and E. Zaccarelli. Modeling equilibrium clusters in lysozyme solutions. *EPL*, 77(4), 2007.

[35] J. H. Carra, E. C. Murphy, and P. L. Privalov. Thermodynamic effects of mutations on the denaturation of t4 lysozyme. *Biophys. J.*, 71:1994–2001, 1996.

[36] R. Carrotta, J. Barthès, A. Longo, V. Martorana, M. Manno, G. Portale, and P. L. San Biagio. Large size fibrillar bundles of the alzheimer amyloid β-protein. *Eur. Biophys. J.*, 36:701–709, 2007.

[37] R. Carrotta, M. Di Carlo, M. Manno, G. Montana, P. Picone, D. Romancino, and P.L. San Biagio. Toxicity of recombinant beta-amyloid prefibrillar oligomers on the morphogenesis of the sea urchin paracentrotus lividus. *FASEB J.*, 20(11):1916, SEP 2006.

[38] R. Carrotta, M. Manno, F. M. Giordano, A. Longo, G. Portale, V. Martorana, and P. L. San Biagio. Protein stability modulated by a conformational effector: effects of trifluoroethanol on bovine serum albumin. *Phys. Chem. Chem. Phys.*, 11:4007–4018, 2009.

[39] B. Caughey and P.T. Lansbury. Protofibrils, pores, fibrils, and neurodegeneration: Separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.*, 26:267–298, 2003.

[40] S. Cavagnero, D.A. Debe, Z.H. Zhou, M.W.W. Adams, and S.I. Chan. Kinetic role of electrostatic interactions in the unfolding of hyperthermophilic and mesophilic rubredoxins. *Biochem.*, 37(10):3369–3376, 1998.

[41] A.S. Celia and D. Volker. The role of protein-solvent interactions in protein unfolding. *Curr. Op. Biotech.*, 7(4):428–432, 1996.

[42] D. Chandler. Interfaces and the driving force of hydrophobic assembly. *NATURE*, 437(7059):640–647, SEP 29 2005.

[43] F. Chiti and C. M. Dobson. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.*, 75:333–366, 2006.

[44] F. Chiti, N. Taddei, P. Webster, D. Hamada, T. Fiaschi, G. Ramponi, and C. M. Dobson. Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nature Struc. Biol.*, 6:380–387, 1999.

[45] A.H. Clark, G.M. Kavanagh, and S.B. Ross-Murphy. Globular protein gelation - theory and experiment. *Food Hydrocoll.*, 15(4-6):383–400, 2001.

[46] P. Claudy, J.M. Létoffé, A. Bayol, M.C. Bonnet, and J.C. Maurizot. Denaturation versus ph of lysozyme and biosynthetic human growth hormone by differential scanning calorimetry and circular dichroism: a comparative study. *Thermochimica Acta*, 207:227–237, 1992.

[47] K.A. Conway, J.D. Harper, and P.T. Lansbury. Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset parkinson disease. *Nature*, 4(11):1318–1320, 1998.

[48] K.A. Conway, S.J. Lee, J.C. Rochet, T.T. Ding, R.E. Williamson, and P.T. Lansbury. Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset parkinson's disease: Implications for pathogenesis and therapy. *PNAS USA*, 97(2):571–576, 2000.

[49] L. Cordone, M. Ferrand, E. Vitrano, and G. Zaccai. Harmonic behavior of trehalose-coated carbon-monoxy-myoglobin at high temperature. *Biophys. J.*, 76:1043–1047, 1999.

[50] A. Cupane, D. Giacomazza, and L. Cordone. Kinetics of thermal denaturation of met-hemoglobin in perturbed solvent: Relevance of bulk-electrostatic and hydrophobic interactions. *Biopolymers*, 21:1081–1092, 1982.

[51] K.A. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990.

[52] Ken A. Dill and Sarina Bromberg. *Molecular Driving Forces: Statistical Thermodynamics in Chemistry & Biology*. Garland Science, 1 edition, September 2002.

[53] T.T. Ding, S.J. Lee, J.C. Rochet, and P.T. Lansbury. Annular alpha-synuclein protofibrils are produced when spherical protofibrils are incubated in solution or bound to brain-derived membranes. *Biochemistry*, 41(32):10209–10217, 2002.

[54] N. M. Dixit and Charles F. Zukoski. Competition between crystallization and gelation: A local description. *Physical Review E*, 67:061501, 2003.

[55] C. M. Dobson. Protein folding and misfolding. *Nature*, 426:884–890, 2003.

[56] C.M. Dobson. Principles of protein folding, misfolding and aggregation. *Seminars in Cell and Developmental Biology*, 15:3–16, 2004.

[57] C.M. Dobson, P.A. Evans, and S.E. Radford. Understanding how proteins fold: the lysozyme story so far. *Trends Biochem. Sci.*, 19:31–37, 1994.

[58] P. Doherty and G. B. Benedek. The effect of electric charge on the diffusion of macromolecules. *J. Chem. Phys.*, 61:5426–5434, 1974.

[59] B.N. Dominy, D. Perl, F.X. Schmid, and C.L. Brooks. The effects of ionic strength on protein stability: The cold shock protein family. *J. Mol. Biol.*, 319(2):541–554, 2002.

[60] A. Dong, J. Matsuura, M. C. Manning, and J. F. Carpenter. Intermolecular β-sheet results from trifluoroethanol-induced nonnative α-helical structure in β-sheet predominant proteins: Infrared and circular dichroism spectroscopic study. *Arch. Biochem. Biophys.*, 335:275–281, 1998.

[61] K.F. DuBay, A.P. Pawar, F. Chiti, J. Zurdo, C.M. Dobson, and M. Vendruscolo. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.*, 341:1317–1326, 2004.

[62] A. Ducruix, J.P. Guilloteau, M. Riés-Kautt, and Tardieu A. Protein interactions as seen by solution x-ray scattering prior to crystallogenesis. *J. Cryst. Growth*, 168(1-4):28–39, 1996.

[63] L. Dunn, C. Carswell, and S. Hagen. Detection of amyloid fibril formation in hen egg-white lysozyme using dynamic light scattering. *Research Experiences for Undergraduates*, unpublished, Summer 2000.

[64] W. Dzwolak, R. Jansen, V. Smirnovas, A. Loksztejn, S. Porowski, and R. Winter. Template-controlled conformational patterns of insulin fibrillar self-assembly reflect history of solvation of the amyloid nuclei. *Phys. Chem. Chem. Phys.*, 7:1349–1351, 2005.

[65] S. Raccosta et al. Lysozyme aggregation induced by thermal unfolding. manuscript in preparation.

[66] S. Raccosta et al. Lysozyme fibrillization and oligomerization driven by competing interactions. manuscript in preparation.

[67] M. Fändrich, V. Forge, K. Buder, M. Kittler, C.M. Dobson, and S. Diekmann. Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments. *PNAS USA*, 100(26):15463–15468, 2003.

[68] G. D. Fasman. *Circular Dichroism and the Conformational Analysis of Biomolecules*. Plenum Press, New York, 1996.

[69] N.L. Fawzi, V. Chubukov, L.A. Clark, S. Brown, and T. Head-Gordon. Influence of denatured and intermediate states of folding on protein aggregation. *Prot. Sci.*, 14:993–1003, 2005.

[70] F. Ferri. Use of a charge coupled device camera for low-angle elastic light scattering. *Rev. Sci. Instrum.*, 68:2265–2274, 1997.

[71] F.A. Ferrone, J. Hofrichter, H.R. Sunshine, and W.A. Eaton. Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. *Biophys. J.*, 32:361–380, 1980.

[72] J. D. Ferry. *Viscoelastic Properties of Polymers*. Wiley, New York, NY, 1980.

[73] S. Finet, F. Skouri-Panet, M. Casselyn, F. Bonneté, and A. Tardieu. The hofmeister effect as seen by saxs in protein solutions. *Current Opinion in Colloid & Interface Science*, 9(1-2):112–116, 2004.

[74] P. J. Flory. *Principles of polymer chemistry*. Cornell University Press, Ithaca, 1953.

[75] H. Flyvbjerg, T.E. Holy, and S. Leibler. Microtubule dynamics: Caps, catastrophes, and coupled hydrolysis. *Phys. Rev. E*, 54(5):5538–5560, 1996.

[76] V. Foderá. *Insulin fibril formation: growth mechanisms and fiber morphologies*. PhD thesis, Dept. Physical and Astronomical Sciences, 2009.

[77] E. Foegeding. Food biophysics of protein gels: A challenge of nano and macroscopic proportions. *Food Biophys.*, 1:41–50, 2006.

[78] E. Frare, M.F. Mossuto, P. Polverino de Laureto, S. Tolin, L. Menzer, M. Dumoulin, C.M. Dobson, and Fontana A. Characterization of oligomeric species on the aggregation pathway of human lysozyme. *J. Mol. Biol.*, 387(1):17–27, 2009.

[79] H. Frauenfelder, F. Parak, and R. D. Young. Conformational substates in proteins. *Ann. Rev. Biophys. Biophys. Chem.*, 17:451–479, 1988.

[80] H. Frauenfelder, S.G. Sligar, and P.G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603, 1991.

[81] E. Freire and E.L. Biltonen. *Biopolymers*, 17:463–479, 1978.

[82] B.J. Frisken. Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Applied Optics*, 40(24):4087–4091, 2001.

[83] C. Frontali, E. Dore, A. Ferrauto, E. Gratton, A. Bettini, M. R. Pozzan, and E. Valdevit. An absolute method for the determination of the persistence length of native dna from electron micrographs. *Biopolymers*, 18:1353–1373, 1979.

[84] J.M. Garcia-Ruiz. Nucleation of protein crystals. *J. Struct. Biol.*, 142(1):22–31, 2003. Macromolecular crystallization in the structural genomics era.

[85] K. Gast, D. Zirwer, M. Muller-Frohne, and G. Damaschun. Trifluoroethanol-induced conformational transitions of proteins: Insights gained from the differences between α-lactalbumin and ribonuclease a. *Protein Sci.*, 8:625–634, 1999.

[86] S. Gavryushov and P. Zielenkiewicz. Electrostatic potential and free energy of proteins: A comparison of the poisson–boltzmann and the bogolyubov–born–green–yvon equations. *J. Phys. Chem. B*, 101(50):10903–10909, 1997.

[87] A. George and W.W. Wilson. Predicting protein crystallization from a dilute solution property. *Acta Crystallogr. D*, 50(4):361–365, 1994.

[88] A.L. Gharibyan, V. Zamotin, K. Yanamandra, O.S. Moskaleva, B.A. Margulis, I.A. Kostanyan, and L.A. Morozova-Roche. Lysozyme amyloid oligomers and fibrils induce cellular death via different apoptotic/necrotic pathways. *J. Mol. Biol.*, 365(5):1337–1349, 2007.

[89] L. Giehm. *Exploring the fibrillation landscape of α-synuclein*. PhD thesis, Dept. of Biotechnology, Chemistry and Environmental Engeneering. University of Aalborg, 2009.

[90] B. Gillespie and K.W. Plaxco. Using protein folding rate to test protein folding theories. *Annu. Rev. Biochem.*, 73(1):837–859, 2004.

[91] S. Goda, K. Takano, Y. Yamagata, R. Nagata, H. Akutsu, S. Maki, K. Namba, and K. Yutani. Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Prot. Sci.*, 9(2):369–375, FEB 2000.

[92] W.S. Gosal and S.B. Ross-Murphy. Globular protein gelation. *Curr. Opin. Colloid Interface Sci.*, 5:188–194, 2000.

[93] N. Greenfield and G.D. Fasman. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry*, 8:4108–4116, 1969.

[94] P. Grochowski and J. Trylska. Continuum molecular electrostatics, salt effects, and counterion binding–a review of the poisson-boltzmann theory and its modifications. *Biopolymers*, 89:93–113, 2008.

[95] M. Groenning, M. Norrman, J.M. Flink, M. van de Weert, J.T. Bukrinsky, G. Schluckebier, and S. Frokjaer. Binding mode of thioflavin t in insulin amyloid fibrils. *J. Struct. Biol.*, 159:483–497, 2007.

[96] M. Groenning, L. Olsen, M. van de Weert, J.M. Flink, S. Frokjaer, and F.S. Jørgensen. Study on the binding of thioflavin t to β-sheet-rich and non-β-sheet cavities. *J. Struct. Biol.*, 158:358–369, 2007.

[97] S. Grudzielanek, R. Jansen, and R. Winter. Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. *J. Mol. Biol.*, 351:879–894, 2005.

[98] C. Guáqueta, L.K. Sanders, G.C.L. Wong, and E. Luijten. The effect of salt on self-assembled actin-lysozyme complexes. *Biophys. J.*, 90, 2006.

[99] A.K. Gupta, C. Coelho, and P.M. Adler. Influence of the stern layer on electrokinetic phenomena in porous media. *J. Coll. Int. Sci.*, 316(1):140–159, 2007.

[100] C. Haas and J. Drenth. Understanding protein crystallization on the basis of the phase diagram. *J. Cryst. Growth*, 196:388–394(7), 15 January 1999.

[101] P. Haezebrouck, M. Joniau, H. Van Dael, S. D. Hooke, N. D. Woodruff, and C. M. Dobson. An equilibrium partially folded state of human lysozyme at low ph. *J. MoL Biol.*, 246:382–387, 1995.

[102] D. Hamada, F. Chiti, J. I. Guijarro, M. Kataoka, N. Taddei, and C. M. Dobson. Evidence concerning rate-limiting steps in protein folding from the effects of trifluoroethanol. *Nature Struc. Biol.*, 7:58–61, 2000.

[103] H. Hamada, T. Arakawa, and K. Shiraki. Effect of additives on protein aggregation. *Curr. Pharm. Biotech.*, 10:400–407, 2009.

[104] Jean Pierre Hansen and Ian R. McDonald. *Theory of simple liquids.* Academic Press, London, New York, San Diego, 1986.

[105] J. Hardy and D.J. Selkoe. The amyloid hypothesis of alzheimer's disease: Progress and problems on the road to therapeutics. *Science*, 297(5580):353–356, 2002.

[106] F.U. Hartl and M. Hayer-Hartl. Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.*, 16:574–581, 2009.

[107] S. Hayakawa and R. Nakamura. Optimization approaches to thermally induced egg-white lysozyme gel. *Agric. Biol. Chem.*, 50:2039–2046, 1986.

[108] J. Jr Hermans and H.A. Scheraga. Structural studies of ribonuclease. v. reversible change of configuration 1-3. *Journal of the American Chemical Society*, 83:3283–3292, 1961.

[109] M. Hilge, G. Siegal, G.W. Vuister, P. Guntert, S.M. Gloor, and J.P. Abrahams. Atp-induced conformational changes of the nucleotide-binding domain of na,k-atpase. *Nat. Struct. Mol. Biol.*, 10:468–474, 2003.

[110] S.E. Hill, J. Robinson, G. Matthews, and M. Muschol. Amyloid protofibrils of lysozyme nucleate and grow via oligomer fusion. *Biophys. J.*, 96(9):3781–3790, 2009.

[111] M.A.M. Hoffmann, J.C. van Miltenburg, J.P. van der Eerden, P.J.J.M. van Mil, and C.G. de Kruif. Isothermal and scanning calorimetry measurements on β-lactoglobulin. *J. Phys. Chem. B*, 101:6988–6994, 1997.

[112] S. Honda, H. Uedaira, F. Vonderviszt, S. Kidokoro, and K. Namba. Folding energetics of a multidomain protein, flagellin. *J. Mol. Biol.*, 293:719–732, 1999.

[113] D.-P. Hong, M. Hoshino, R. Kuboi, and Y. Goto. Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *J. Am. Chem. Soc.*, 121:8427–8433, 1999.

[114] D.P. Hong, A.L. Fink, and V.N. Uversky. Structural characteristics of alpha-synuclein oligomers stabilized by the flavonoid baicalein. *J. Mol. Biol.*, 383(1):214–223, OCT 31 2008.

[115] M. Hoshino, Y. Hagihara, D. Hamada, M. Kataoka, and Y. Goto. Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle x-ray scattering. *FEBS Lett.*, 416:72–76, 1997.

[116] Q.R. Huang, P.L. Dubin, C.N. Moorefield, and G.R. Newkome. Counterion binding on charged spheres: Effect of ph and ionic strength on the mobility of carboxyl-terminated dendrimers. *J. Phys. Chem. B*, 104(5):898–904, 2000.

[117] H.R. Ibrahim, U. Thomas, and A. Pellegrini. A helix-loop-helix peptide at the upper lip of the active site cleft of lysozyme confers potent antimicrobial activity with membrane permeabilization action. *J. Biol. Chem.*, 276(47):43767–43774, 2001.

[118] A. Isihara. Determination of molecular shape by osmotic measurement. *J. Chem. Phys.*, 18:1446–1449, 1950.

[119] J. Janis, J. Rouvinen, P. Vainiotalo, O. Turunen, and V.L. Shnyrov. Irreversible thermal denaturation of trichoderma reesei endo-1,4-beta-xylanase ii and its three disulfide mutants characterized by differential scanning calorimetry. *Int. J. Biol. Macromol.*, 42(1):75–80, JAN 1 2008.

[120] N. Javid, K. Vogtt, C. Krywka, M. Tolan, and R. Winter. Protein-protein interactions in complex cosolvent solutions. *ChemPhysChem*, 8:679–689, 2007.

[121] J. L. Jimenez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil. The protofilament structure of insulin amyloid fibrils. *Protein. Sci.*, 9:1960–1967, 2000.

[122] Wyman J. Jr. Linked functions and reciprocal effects in hemoglobin: A second look. volume 19 of *Advances in Protein Chemistry*, pages 223–286. Academic Press, 1964.

[123] A. Kato, K. Fujimoto, N. Matsudomi, and K. Kobayashi. Protein flexibility and functional properties of heat-denatured ovalbumin and lysozyme. *Agric. Biol. Chem.*, 50:417–420, 1986.

[124] N. Kishore and B. Sabulal. Thermal unfolding of hen egg-white lysozyme in the presence of 4-chlorobutan-1-ol. *Pure and Appl. Chem.*, 70(3):665–670, 1998.

[125] P.V. Konarev, M.V. Petoukhov, V.V. Volkov, and D.I. Svergun. Atsas 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, 39(Part 2):277–286, 2006.

[126] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, and D.I. Svergun. Primus: a windows pc-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, 36(Part 5):1277–1282, 2003.

[127] D.E. Koppel. Analysis of macromolecular polydispersity in intensity correlation spectroscopy: The method of cumulants. *J. Chem. Phys.*, 57(11):4814–4820, 1972.

[128] M. R. H. Krebs, E. H. C. Bromley, and A. M. Donald. The binding of thioflavin-t to amyloid fibrils: localisation and implications. *J. Struct. Biol.*, 149:30–37, 2005.

[129] M. R. H. Krebs, G. L. Devlin, and A. M. Donald. Protein particulates: Another generic form of protein aggregation? *Biophys. J.*, 92:1336–1342, 2007.

[130] M. R. H. Krebs, D. K. Wilkins, E. W. Chung, M. C. Pitkeathly, A. K. Chamberlain, J. Zurdo, C. V. Robinson, and C. M. Dobson. Formation and Seeding of Amyloid Fibrils from Wild-type Hen Lysozyme and a Peptide Fragment from the β-Domain. *J. Mol. Biol.*, 300:541–549, 2000.

[131] B.I. Kurganov, A.E. Lyubarev, J.M. Sanchez-Ruiz, and V.L. Shnyrov. Analysis of differential scanning calorimetry data for proteins - criteria of validity of one-step mechanism of irreversible protein denaturation. *Biophys. Chem.*, 69(2-3):125–135, 1997.

[132] B. Lai, A. Cao, and L. Lai. Organic cosolvents and hen egg white lysozyme folding. *Biophys. Biochem. Acta*, 1543:115–122, 2000.

[133] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Kluwer Academic/Plenum Publishers, New York, 1999.

[134] L.D. Landau and E.M. Lifshitz. *Statistical Physics*, volume 5 of *Course of Theoretical Physics*. Pergamon Press, third edition, 1980.

[135] A.E. Langkilde and B. Vestergaard. Methods for structural characterization of prefibrillar intermediates and amyloid fibrils. *FEBS Letters*, 583(16):2600–2609, 2009.

[136] H.A. Lashuel, D. Hartley, B.M. Petre, T. Walz, and P.T. Lansbury. Neurodegenerative disease: Amyloid pores from pathogenic mutations. *Nature*, 418(11):291, 2002.

[137] H.A. Lashuel, B.M. Petre, J. Wall, M. Simon, R.J. Nowak, T. Walz, and P.T. Lansbury. alpha-synuclein, especially the parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. *J. Mol. Biol.*, 322(5):1089–1102, 2002.

[138] E.E. Lattman and G.D. Rose. Protein folding–what's the question? *PNAS USA*, 90(2):439–441, 1993.

[139] H. Levine III. Thioflavine t interaction with synthetic alzheimer's disease beta-amyloid peptides: Detection of amyloid aggregation in solution. *Prot. Sci.*, 2:404–410, 1993.

[140] H. Li, A.D. Robertson, and J.H. Jensen. Very fast empirical prediction and rationalization of protein pka values. *Proteins*, 61:704–721, 2005.

[141] W. Liu, D. Bratko, J. M. Prausnitz, and H. W. Blanch. Effect of alcohols on aqueous lysozyme-lysozyme interactions from static light-scattering measurements. *Biophys. Chem.*, 107:289–298, 2004.

[142] W. Liu, J. M. Prausnitz, and H. W. Blanch. Effect of alcohols on aqueous lysozyme-lysozyme interactions from static light-scattering measurements. *Biomacromolecules*, 105:1818–1823, 2004.

[143] J.J. Lopez-Garcia, M.J. Aranda-Rascon, and J. Horno. Electrical double layer around a spherical colloid particle: The excluded volume effect. *J. Coll. Int. Sci.*, 316(1):196–201, 2007.

[144] R. Lowe, D.L. Pountney, P.H. Jensen, W.P. Gai, and N.H. Voelcker. Calcium(ii) selectively induces alpha-synuclein annular oligomers via interaction with the c-terminal domain. *Prot. Sci.*, 13(12):3245–3252, 2004.

[145] H. Lu, M. Buck, S. E. Radford, and C. M. Dobson. Acceleration of the folding of hen lysozyme by trifluoroethanol. *J. Mol. Biol.*, 265:112–117, 1997.

[146] M. Lund and B. Jonsson. On the charge regulation of proteins. *Biochemistry*, 44(15):5722–5727, 2005.

[147] Y. Luo and R. L. Baldwin. Trifluoroethanol stabilizes the ph 4 folding intermediate of sperm whale apomyoglobin. *J. Mol. Biol.*, 279:49–57, 1998.

[148] A.E. Lyubarev, B.I. Kurganov, A.A. Burlakova, and V.N. Orlov. Irreversible thermal denaturation of uridine phosphorylase from escherichia coli k-12. *Biophys. Chem.*, 70(3):247–257, 1998.

[149] G.S. Manning. Limiting laws and cunterion condensation in polyelectrolyte solutions. 1 colligative properties. *J. Chem. Phys.*, 51:924–933, 1969.

[150] G.S. Manning. Counterion condensation on charged spheres, cylinders, and planes. *J. Phys. Chem. B*, 111(29):8554–8559, 2007. 6th International Symposium on Polyelectrolytes, Dresden, GERMANY, SEP, 2006.

[151] M. Manno, E. F. Craparo, A. Podestà, D. Bulone, R. Carrotta, V. Martorana, G. Tiana, and P. L. San Biagio. Kinetics of different processes in human insulin amyloid formation. *J. Mol. Biol.*, 366:258–274, 2007.

[152] Y.J. Mao, X.R. Sheng, and X.M. Pan. The effects of nacl concentration and ph on the stability of hyperthermophilic protein ssh10b. *BMC Biochemistry*, 8(1):28, 2007.

[153] S. R. Martin, V. Esposito, P. De Los Rios, A. Pastore, and P. A. Temussi. Cold denaturation of yeast frataxin offers the clue to understand the effect of alcohols on protein stability. *J. Am. Chem. Soc.*, 130:9963–9970, 2008.

[154] A.A. Maskevich, V.I. Stsiapura, V.A. Kuzmitsky, I.M. Kuznetsova, O.I. Povarova, V.N. Uversky, and K.K. Turoverov. Spectral properties of thioflavin t in solvents with different dielectric properties and in a fibril-incorporated form. *J. Proteome Res.*, 6:1392–1401, 2007.

[155] A. Matagne, M. Jamin, E. W. Chung, C. V. Robinson, S. E. Radford, and C. M. Dobson. Thermal unfolding of an intermediate is associated with non-arrhenius kinetics in the folding of hen lysozyme. *J. Mol. Biol.*, 297:193–210, 2000.

[156] H. Mattoussi, S. O'Donohue, and F.E. Karasz. Polyion conformation and second virial coefficient dependences on the ionic strength for flexible polyelectrolyte solutions. *Macromol.*, 25(2):743–749, 1992.

[157] A. Michnik, Z. Drzazga, A. Kluczewska, and K. Michalik. Differential scanning microcalorimetry study of the thermal denaturation of haemoglobin. *Biophys. Chem.*, 118:93–101, 2005.

[158] V. Militello, V. Vetri, and M. Leone. Conformational changes involved in thermal aggregation processes of bovine serum albumin. *Biophys. Chem.*, 105(1):133–141, 2003.

[159] P. Montagne, M.L. Cuilliere, C. Mole, M.C. Bene, and G. Faure. Microparticle-enhanced nephelometric immunoassay of lysozyme in milk and other human body fluids. *Clin. Chem.*, 44(8):1610–1615, 1998.

[160] Y.U. Moon, C.O. Anderson, H.W. Blanch, and J.M. Prausnitz. Osmotic pressures and second virial coefficients for aqueous saline solutions of lysozyme. *Fluid Phase Equilib.*, 168(2):229–239, FEB 29 2000.

[161] L.A. Morozova-Roche, V. Zamotin, M. Malisauskas, A. Öhman, R. Chertkova, M.A. Lavrikova, I.A. Kostanyan, D.A. Dolgikh, and M.P. Kirpichnikov. Fibrillation of carrier protein albebetin and its biologically active constructs. multiple oligomeric intermediates and pathways. *Biochemistry*, 43:9610–9619, 2004.

[162] L.A. Morozova-Roche, J. Zurdo, A. Spencer, W. Noppe, V. Receveur, D.B. Archer, M. Joniau, and C.M. Dobson. Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J. Struct. Biol.*, 130:339–351, 2000.

[163] V. Munoz and J.M. Sanchez-Ruiz. Exploring protein-folding ensembles: A variable-barrier model for the analysis of equilibrium unfolding experiments. *PNAS USA*, 101(51):17646–17651, DEC 21 2004.

[164] J. K. Myers, C. N. Pace, and J. M. Scholtz. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.*, 4:2138–2148, 1995.

[165] H. Naiki, K. Higuchi, M. Hosokawa, and T. Takeda. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine t. *Analyt. Biochem.*, 177:244–249, 1989.

[166] B.L. Neal, D. Asthagiri, and A.M. Lenhoff. Molecular origins of osmotic second virial coefficients of proteins. *Biophys. J.*, 75:2469–2477, 1998.

[167] D. L. Nelson and M. M. Cox. *I principi di biochimica di Lehninger*. Zanichelli, Bologna, 2002.

[168] J.A. Ng, T. Vora, V. Krishnamurthy, and S.H. Chung. Estimating the dielectric constant of the channel protein and pore. *Eur. Biophys. J. Biophys. Lett.*, 37(2):213–222, 2008.

[169] T. Nicolai and D. Durand. Protein aggregation and gel formation studied with scattering methods and computer simulations. *Curr. Op. Coll. & Int. Sci.*, 12(1):23–28, 2007.

[170] T. Nicolai, D. Durand, and J.-C. Gimel. *In: Light scattering: principles and development, (ed. W. Brown)*. Clarendon Press, Oxford, 1996.

[171] D. Nicoli and G. Benedek. Study of thermal denaturation of lysozyme and other globular proteins by light-scattering spectroscopy. *Biopolymers*, 15:2421–2437, 1976.

[172] L. Nielsen, R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V. N. Uversky, and A. L. Fink. Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry*, 40:6036–6046, 2001.

[173] D. Nohara, A. Mizutani, and T. Sakai. Kinetic study on thermal denaturation of hen egg-white lysozyme involving precipitation. *J. Biosci. Bioeng.*, 87:199–205, 1999.

[174] W. Ostwald. Studies upon the forming and changing solid bodies. *Physikalische Chemie*, 22:289–330, 1897.

[175] M. Panouille, T. Nicolai, and D. Durand. Heat induced aggregation and gelation of casein submicelles. *International Dairy Journal*, 14:297–303, 2004.

[176] A.S. Parmar and M. Muschol. Lysozyme as diffusion tracer for measuring aqueous solution viscosity. *J. Coll. Int. Sci.*, 339(1):243–248, 2009.

[177] R. M. Parodi, E. Bianchi, and A. Ciferri. Thermodynamics of unfolding of lysozyme in aqueous alcohol solutions. *J. Biol. Chem.*, 248:4047–4051, 1973.

[178] L. Pauling, H.A. Itano, S.J. Singer, and I.C. Wells. Sickle cell anemia, a molecular disease. *Science*, 110(2865):543–548, 1949.

[179] J. S. Pedersen. Small-angle scattering from precipitates: Analysis by use of a polydisperse hard-sphere model. *Phys. Rev. B*, 47:657–665, 1993.

[180] J.S. Pedersen, D. Dikov, J.L. Flink, H.A. Hjuler, G. Christiansen, and D.E. Otzen. The changing face of glucagon fibrillation: Structural polymorphism and conformational imprinting. *J. Mol. Biol.*, 355:501–523, 2006.

[181] A.T. Petkova, R.D. Leapman, Z. Guo, W.M. Yau, M.P. Mattson, and R. Tycko. Self-propagating, molecular-level polymorphism in alzheimer's β-amyloid fibrils. *Science*, 307(5707):262–265, 2005.

[182] D.N. Petsev, X. Wu, O. Galkin, and P.G. Vekilov. Thermodynamic functions of concentrated protein solutions from phase equilibria. *J. Phys. Chem. B*, 107(16):3921–3926, 2003.

[183] R. Piazza and S. Iacopini. Transient clustering in a protein solution. *Eur. Phys. J. E*, 7:45–48, 2002.

[184] Roberto Piazza, Véronique Peyre, and Vittorio Degiorgio. -sticky hard sphere-model of proteins near crystallization: A test based on the osmotic compressibility of lysozyme solutions. *Phys. Rev. E*, 58:R2733, 1998.

[185] P. Picone, R. Carrotta, G. Montana, M.R. Nobile, P.L. San Biagio, and M. Di Carlo. A beta oligomers and fibrillar aggregates induce different apoptotic pathways in lan5 neuroblastoma cell cultures. *Biophys. J.*, 96(10):4200–4211, 2009.

[186] A. Podestà, G. Tiana, P. Milani, and M. Manno. Early events in insulin fibrillization studied by time-lapse atomic force microscopy. *Biophys. J.*, 90:589–597, 2006.

[187] M.H. Polymeropoulos, C. Lavedan, E. Leroy, S.E. Ide, A. Dehejia, A. Dutra, B. Pike, H. Root, J. Rubenstein, R. Boyer, E.S. Stenroos, S. Chandrasekharappa, A. Athanassiadou, T. Papapetropoulos, W.G. Johnson, A.M. Lazzarini, R.C. Duvoisin, G. Di Iorio, L.I. Golbe, and R.L. Nussbaum. Mutation in the α-synuclein gene identified in families with parkinson's disease. *Science*, 276(5321):2045–2047, 1997.

[188] W.C.K. Poon, S.U. Egelhaaf, P.A. Beales, A. Salonen, and L. Sawyer. Protein crystallization: scaling of charge and salt concentration in lysozyme solutions. *J. Phys.: Cond. Matt.*, 12:L569–L574, 2000.

[189] J. Povey, C. M. Smales, S. J. Hassard, and M. J. Howard. Comparison of the effects of 2,2,2-trifluoroethanol on peptide and protein structure and function. *J. Struct. Biol.*, 157:329–338, 2007.

[190] E.T. Powers and D.L. Powers. The kinetics of nucleated polymerizations at high concentrations: Amyloid fibril formation near and above the "supercritical concentration". *Biophys. J.*, 91(1):122–132, 2006.

[191] J. M. Prausnitz. Molecular thermodynamics for some applications in biotechnology. *Pure Appl. Chem.*, 75:859–873, 2003.

[192] P. L. Privalov. Intermediate states in protein folding. *J. Mol. Biol.*, 258:707–725, 1996.

[193] P.L. Privalov. Stability of proteins. small globular proteins. *Adv. Prot. Chem.*, 33:167–241, 1979.

[194] P.L. Privalov. Stability of proteins: Proteins which do not present a single cooperative system. volume 35 of *Advances in Protein Chemistry*, pages 1–104. Academic Press, 1982.

[195] P.L. Privalov. Thermodynamic problems of protein structure. *Annu. Rev. Biophys. Biophys. Chem.*, 18(1):47–69, 1989.

[196] P.L. Privalov and N.N. Khechinashvili. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.*, 86(3):665–684, 1974.

[197] P.L. Privalov and S.A. Potekhin. Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.*, 131:4–51, 1986.

[198] S.W. Provencher. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Computer Phys. Communicat.*, 27(3):213–227, 1982.

[199] S.W. Provencher. Contin: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Computer Phys. Communicat.*, 27(3):229–242, 1982.

[200] S.W. Provencher and J. Gloeckner. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, 20(1):33–37, 1981.

[201] S.B. Prusiner. Prions. *PNAS*, 95:13363–13383, 1998.

[202] A.M. Puertas, C. De Michele, F. Sciortino, P. Tartaglia, and E. Zaccarelli. Viscoelasticity and stokes-einstein relation in repulsive and attractive colloidal glasses. *J. Chem. Phys.*, 127(14):144906, 2007.

[203] P.N. Pusey. *Liquides, cristallisation et transition vitreuse. Part II. Liquids, freezing and glass transition. By J. P. Hansen. D. Levesque and J. Zinn-Justin. North-Holland, 1989. pp. 995.* North-Holland, Amsterdam, 1989.

[204] S. Raccosta, M. Manno, D. Bulone, D. Giacomazza, V. Militello, V. Martorana, and P.L. Biagio. Irreversible gelation of thermally unfolded proteins: structural and mechanical properties of lysozyme aggregates. *Eur. Biophys. J.*, 39(6):1007–17, 2010.

[205] S. E. Radford, C. M. Dobson, and P. A. Evans. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature (London)*, 358:302–307, 1992.

[206] R. Rajan and P. Balaram. A model for the interaction of trifluoroethanol with peptides and proteins. *Int. J. Peptide Protein Res.*, 48:328–336, 1996.

[207] H. Reiersen and A. R. Rees. Trifluoroethanol may form a solvent matrix for assisted hydrophobic interactions between peptide side chains. *Protein Engineering*, 13:739–743, 2000.

[208] R.L. Remmele, J.Z. van Enk, V. Dharmavaram, D. Balaban, M. Durst, A. Shoshitaishvili, and Hugh R. Scan-rate-dependent melting transitions of interleukin-1 receptor (type ii): Elucidation of meaningful thermodynamic and kinetic parameters of aggregation acquired from dsc simulations. *Journal of the American Chemical Society*, 127:8328–8339, 2005.

[209] K.F. Riley, M.P. Hobson, and S.J. Bence. *Mathematical methods for physics and engineering.* Cambridge University Press, 3 edition, March 2006.

[210] C.J. Roberts. Non-native protein aggregation kinetics. *Biotech. Bioeng.*, 98(5):927–938, 2007.

[211] D. Roccatano, G. Colombo, M. Fioroni, and A. E. Mark. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proc. Natl. Acad. Sci. USA*, 99:12179–12184, 2002.

[212] T.W. Rosch and Errington J.R. Investigation of the phase behavior of an embedded charge protein model through molecular simulation. *J. Phys. Chem. B*, 111:12591–12598, 2007.

[213] D. Rosenbaum, P. C. Zamora, and C. F. Zukoski. Phase behavior of small attractive colloidal particles. *Physical Review Letters*, 76(1):150–153, 1996.

[214] D. F. Rosenbaum, Amit Kulkarni, S. Ramakrishnan, and C. F. Zukoski. Protein interactions and phase behavior: Sensitivity to the form of the pair potential. *The Journal of Chemical Physics*, 111(21):9882–9890, 1999.

[215] D. M. Rothwarf and H. A. Scheraga. Role of non-native aromatic and hydrophobic interactions in the folding of hen egg white lysozyme. *Biochemistry*, 35:13797–13807, 1996.

[216] P.E. Jr Rouse. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *J. Chem. Phys.*, 21(7):1272–1280, 1953.

[217] P. Salma, C. Chhatbar, and S. Seshadri. Intrinsically unstructured proteins: Potential targets for drug discovery. *Am. J. Infect. Diseases*, 5:126–134, 2009.

[218] G. Salvetti, E. Tombari, and L. Mikheeva. The endothermic effects during denaturation of lysozyme by temperature modulated calorimetry and an intermediate reaction equilibrium. *J. Phys. Chem. B*, 106:6081–6087, 2002.

[219] J.M. Sanchez-Ruiz. Theoretical-analysis of lumry-eyring models in differential scanning calorimetry. *Biophys. J.*, 61(4):921–935, 1992.

[220] J.M. Sánchez-Ruiz, J.L. López-Lacomba, M. Cortijo, and P.L. Mateo. Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. *Biochemistry*, 27:1648–1652, 1988.

[221] J.M. Sanchez-Ruiz and G.I. Makhatadze. To charge or not to charge? *Trends Biotech.*, 19(4):132–135, 2001.

[222] M. Sandal, F. Valle, I. Tessari, S. Mammi, E. Bergantino, F. Musiani, M. Brucale, L. Bubacco, and B. Samori. Conformational equilibria in monomeric alpha-synuclein at the single-molecule level. *PLoS Biol.*, 6(1):99–108, 2008.

[223] K. Sasahara, M. Demura, and K. Nitta. Partially unfolded equilibrium state of hen lysozyme studied by circular dichroism spectroscopy. *Biochemistry*, 39:6475–6482, 2000.

[224] G. Schiró, M. Sclafani, F. Natali, and A. Cupane. Hydration dependent dynamics in sol-gel encapsulated myoglobin. *Eur. Biophys. J.*, 37:543–549, 2008.

[225] M. Schlierf and M. Rief. Temperature softening of a protein in single-molecule experiments. *J. Mol. Biol.*, 354(2):497–503, 2005.

[226] K. S. Schmitz. *An Introduction to Dynamic Light Scattering by Macromolecules*. Academic Press, Inc., San Diego, 1990.

[227] A. Schukla, E. Mylonas, E. Di Cola, S. Finet, P. Timmins, T. Narayanan, and D.I. Svergun. Absence of equilibrium cluster phase in concentrated lysozyme solutions. *PNAS USA*, 105(13):5075–5080, 2008.

[228] F. Sciortino, P. Tartaglia, and E. Zaccarelli. One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion. *J. Phys. Chem. B*, 109(46):21942–21953, 2005.

[229] L.C. Serpell, C.C.F. Blake, and P.E. Fraser. Molecular structure of a fibrillar alzheimer's a β fragment. *Biochemistry*, 39:13269–13275, 2000.

[230] Li Shaoxin, Xing Da, and Li Junfeng. Dynamic light scattering application to study protein interactions in electrolyte solutions. *Journal of Biological Physics*, 30:313–324, 2004.

[231] I. Shin, E. Wachtel, E. Roth, C. Bon, I. Silman, and L. Weiner. Thermal denaturation of bungarus fasciatus acetylcholinesterase: Is aggregation a driving force in protein unfolding? *Prot. Sci.*, 11:2022–2032, 2002.

[232] K. Shiraki, K. Nishikawa, and Y. Goto. Trifluoroethanol-induced stabilization of the α-helical structure of β-lactoglobulin: Implication for non-hierarchical protein folding. *J. Mol. Biol.*, 245:180–194, 1995.

[233] A. Shukla, E. Mylonas, E. Di Cola, S. Finet, P. Timmins, T. Narayanan, and D. Svergun. Absence of equilibrium cluster phase in concentrated lysozyme solutions. *PNAS USA*, 105:5075–5080, 2008.

[234] S. Singh and J. Singh. Effect of polyols on the conformational stability and biological activity of a model protein lysozyme. *AAPS Pharm. Sci. Tech.*, 4:101–109, 2003.

[235] M. Stefani and C. M. Dobson. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.*, 81:678–699, 2003.

[236] A. Stradner, H. Sedgwick, F. Cardinaux, W.C.K. Poon, S.U. Egelhaaf, and P. Schurtenberger. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432:492–495, 2004.

[237] A.M. Streets and S.R. Quake. Ostwald ripening of clusters during protein crystallization. *Phys. Rev. Lett.*, 104(17):178102, 2010.

[238] V.I. Stsiapura and A.A. Maskevich. Thioflavin t as a molecular rotor: Fluorescent properties of thioflavin t in solvents with different viscosity. *J. Phys. Chem. B*, 112:15893–15902, 2008.

[239] V.I. Stsiapura, A.A. Maskevich, V.A. Kuzmitsky, K.K. Turoverov, and I.M. Kuznetsova. Computational study of thioflavin t torsional relaxation in the excited state. *J. Phys. Chem. A*, 111:4829–4835, 2007.

[240] D.I. Svergun. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing (vol 76, pg 2879, 1999). *Biophys. J.*, 77(5):2896, 1999.

[241] D.I. Svergun and M.H.J. Koch. Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, 66(10):1735, 2003.

[242] A. Szilagyi and P. Zavodszky. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure*, 8(5):493–504, 2000.

[243] K. Takano, K. Tsuchimori, Y. Yamagata, and K. Yutani. Contribution of salt bridges near the surface of a protein to the conformational stability,. *Biochemistry*, 39:12375–12381, 2000.

[244] C. Tanford. *Physical Chemistry of Macromolecules*. Wiley, New York, 1961.

[245] C. Tanford. Protein denaturation. c. theoretical models for the mechanism of denaturation. *Adv. Protein Chem.*, 24:1–95, 1970.

[246] C. Tanford and J. Epstein. The physical chemistry of insulin. i. hydrogen ion titration curve of zinc-free insulin. *J. Am. Chem. Soc.*, 76:2163–2169, 1954.

[247] A. Tantos, P. Friedrich, and P. Tompa. Cold stability of intrinsically disordered proteins. *FEBS Letters*, 583(2):465–469, JAN 22 2009.

[248] D. M. Taverna and R. A. Goldstein. Why are proteins marginally stable? *Proteins*, 46:105–109, 2002.

[249] P.A. Temussi, L. Masino, and A. Pastore. From alzheimer to huntington: why is a structural understanding so difficult? *EMBO J.*, 22(3):355–361, 2003.

[250] P.M. Tessier, A.M. Lenhoff, and S.I. Sandler. Rapid measurement of protein osmotic second virial coefficients by self-interaction chromatography. *Biophys. J.*, 82:1620–1631, 2002.

[251] P. D. Thomas and K. A. Dill. Local and nonlocal interactions in globular proteins and mechanisms of alcohol denaturation. *Protein Sci.*, 2:2050–2065, 1993.

[252] S. N. Timasheff. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. USA*, 99:9721–9726, 2002.

[253] T. Ueda, K. Masumoto, R. Ishibashi, T. So, and T. Imoto. Remarkable thermal stability of doubly intramolecularly cross-linked hen lysozyme. *Prot. Eng.*, 13(3):193–196, 2000.

[254] V.N. Uversky. Neuropathology, biochemistry, and biophysics of α-synuclein aggregation. *J. Neurochem.*, 103:17–37, 2007.

[255] I. van der Plancken, A. van Loey, and M.E. Hendrickx. Effect of heat-treatment on the physico-chemical properties of egg white proteins: A kinetic study. *J. Food Eng.*, 75:316–326, 2006.

[256] P.S. Vassar and C.F.A. Culling. Fluorescent stains with special reference to amyloid and connective tissue. *Arch. pathol.*, 68:487, 1959.

[257] G. Velicelebi and J. M. Sturtevant. Thermodynamics of the denaturation of lysozyme in alcohol-water mixtures. *Biochemistry*, 18:1180–1186, 1979.

[258] M. Vendruscolo and G.G. Tartaglia. Towards quantitative predictions in cell biology using chemical properties of proteins. *Mol. BioSyst.*, 4:1170–1175, 2008.

[259] A.W.P. Vermeer and W. Norde. The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein. *Biophys. J.*, 78:394–404, 2000.

[260] B.A. Vernaglia, J. Huang, and E.D. Clark. Guanidine hydrochloride can induce amyloid fibril formation from hen egg-white lysozyme. *Biomacromol.*, 5(4):1362–1370, 2004.

[261] B. Vestergaard, M. Groenning, M. Roessle, J.S. Kastrup, M. van de Weert, J.M. Flink, S. Frokjaer, M. Gajhede, and D.I. Svergun. A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *PLOS Biol.*, 5(5):1089–1097, 2007.

[262] V.V. Volkov and D.I. Svergun. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.*, 36(3):860–864, 2003.

[263] M.J. Volles and P.T. Lansbury. Zeroing in on the pathogenic form of alpha-synuclein and its mechanism of neurotoxicity in parkinson's disease. *Biochemistry*, 42(26):7871–7878, 2003.

[264] E.S. Voropai, M.P. Samtsov, K.N. Kaplevskii, A.A. Maskevich, V.I. Stepuro, O.I. Povarova, I.M. Kuznetsova, K.K. Turoverov, A.L. Fink, and V.N. Uversky. Spectral properties of thioflavin t and its complexes with amyloid fibrils. *Journal of Applied Spectroscopy*, 70:868–874, 2003. 10.1023/B:JAPS.0000016303.37573.7e.

[265] R. Walgers, T. C. Lee, and A. Cammers-Goodwin. An indirect chaotropic mechanism for the stabilization of helix conformation of peptides in aqueous trifluoroethanol and hexafluoro-2-propanol. *J. Am. Chem. Soc.*, 120:5073–5079, 1998.

[266] H.H. Winter and F. Chambon. Analysis of linear viscoelasticity of a crosslinking polymer at the gel point. *J. Rheol.*, 30(2):367–382, 1986.

[267] M. Wogulis, S. Wright, D. Cunningham, T. Chilcote, K. Powell, and R.E. Rydel. Nucleation-dependent polymerization is an essential component of amyloid-mediated neuronal cell death. *J. Neurosci.*, 25(5):1071–1080, 2005.

[268] C.E. Woodward and B. Jonsson. A poisson-boltzmann approximation for strongly interacting macroionic solutions. *J. Phys. Chem.*, 92(7):2000–2007, 1988.

[269] H. Yan, A. Saiani, J.E. Gough, and A.F. Miller. Thermoreversible protein hydrogel as cell scaffold. *Biomacromol.*, 7:2776–2782, 2006.

[270] G.G. Zhadan and V.L. Shnyrov. Differential-scanning-calorimetric study of the irreversible thermal denaturation of 8 kda cytotoxin from the sea anemone radianthus macrodactylus. *Biochem. J.*, 299(3):731–733, 1994.

Acknowledgments

This work was partially supported by the Italian National Research Council through the project *Intermolecular interaction in protein metastable solution*. I want to thank Dr. San Biagio for giving me all the equipment for scientific experiments. For the same reason I want to thank Prof. Cupane for giving me the permission to use the calorimeter in his lab and Prof. Leone, my supervisor, for the permission to use AFM equipment and, especially, for his willingness at any critical moment. In a so large work, like a PhD thesis, people that contributed in one way or another are innumerable. Thus probably I might forget someone and I start to apologize now. I wish to thank all Professors who, during these years, stimulated me intellectually and always increased my desire for knowledge. This thesis work is also the result of several interesting discussions, which were useful to improve the quality of this work. In particular I want to thank Prof. Cupane and Prof. Emanuele for useful discussions (truly productive!). I wish to thank Prof. Compagno for the useful suggestions during the final review of this thesis. I wish to thank Prof. Vestergaard for a wonderful welcome in Denmark, for her precious and irreplaceable assistance in data analysis and for conveying to me the desire not to surrender before the difficulties. I want to thank also all the members of the SAXS group for useful discussions during our very long meetings and for having listened to my complaints about weather in Denmark for six months (but I liked it in June and July!), in particular Annette, Katrine, Malene and Magda. A special thanks goes to Magda, who helped me a lot especially at the beginning with the accomodation. I want also to thank Lise for her scientific availability and the other people of "Fibril group" and in Medicinal Chemistry department. I wish also to thank my friends in Denmark, in particular Claudia, Monica, Fatemeh, Tanya, Elena, Ruzanna and $\cdots$ Olga. Then, an incalculable gratitude goes to Mauro, my co-supervisor, who taught me everything or almost (other researchers at IBF taught me the rest). In particular I wish to thank Enzo, Donatella, Daniela, Rosetta, Mariuccia and Rita. I want to thank my best friends, with whom I spent long days at CNR and in pubs of Palermo (after work!): Paola, Valeria, Daniele, Christian and Tony. Of course a dear embrace goes to other my friends: Massimo, Roberto, Salvatore, $\cdots$ and also my friends on Facebook! (they will understand if I do not name them one by one). As I wrote in the dedication, I want to dedicate this huge work to my mother and my father. They have supported me in every choice and helped me in anyway they could. Probably there are no appropriate words to thank them, so I will leave my heart finds the right words. A special mention goes to my father: he has done so many things for me and now it is arrived the moment that I can do something for him. Finally I want to thank God, who gave me the strength and the constancy to reach this target. Now that also this stage of my life has passed, He is the only person who can drive me towards another adventure.

Samuele Raccosta Palermo, Feb 15^{th} 2011